

Desarrollo de líneas de cacao (*Theobroma cacao* L.) editadas genéticamente que limiten la acumulación de cadmio en sus tejidos: FASE-I

Jose Ángel Rodríguez Bolaño^a, Neyder Alejandro Marín Vélez^a, Diego Fernando Villanueva-Mejía^b

^aPregrado en Biología, Departamento de Ciencias Biológicas, Escuela de Ciencias, Universidad EAFIT, Medellín, Colombia.

^bDepartamento de Ciencias Biológicas, Escuela de Ciencias, Universidad EAFIT, Medellín, Colombia.

Resumen

El cacao (*Theobroma cacao*) es la fuente de materia prima para la elaboración de chocolate y derivados. En las últimas décadas se han encontrado altas concentraciones de cadmio en productos a base de cacao, lo cual se ha visto por distintas organizaciones como una amenaza a la salud humana, la seguridad alimentaria y el futuro de la industria cacaotera. Por tal motivo, es necesario el desarrollo de metodologías en manejo agronómico e ingeniería genética que conduzcan a la reducción de los niveles de este metal pesado en plantas. En el presente estudio se realizó un análisis de metadatos *in silico* que permitiera la identificación de genes implicados en el proceso de absorción de cadmio para postularlos como objeto de mejoramiento genético mediante técnicas de biotecnología moderna de precisión, como la edición genética. Se encontraron 26 proteínas, 88% reportadas en *Arabidopsis thaliana* y 38% en *Oryza sativa*, las cuales pertenecían a 8 familias proteicas. Los miembros restantes de cada familia fueron tenidos en cuenta para los análisis. Así, se evaluaron datos de expresión obtenidos del software Genevestigator® para 44 genes de *A. thaliana* y 37 de *O. sativa*, de los cuales se encontró que, el gen *TcNRAMP5* de *T. cacao*, “ortólogo” de *OsNRAMP5* de *O. sativa*, está estrechamente relacionado al proceso de absorción radicular de Cd, y en consecuencia, se postuló como gen objeto de edición genética en búsqueda de disminuir la concentración de Cd en plantas de cacao. Asimismo, se estructuró una metodología de edición genética basada en el sistema CRISPR/Cas9 mediante transformación genética mediada por *Agrobacterium tumefaciens* y transfección de protoplastos. Además, con esta investigación se estructuró una metodología para establecimiento y generación de plantas *in vitro* mediante embriogénesis somática indirecta y el aislamiento de protoplastos, base fundamental de la plataforma biotecnológica moderna propuesta, que conlleva al mejoramiento genético de precisión en cacao.

Abstract

Cocoa (*Theobroma cacao*) is the primary source of raw material for chocolate production and its derivatives. In current decades, high concentrations of cadmium have been found in cocoa-based products, which has been recognized by several organizations as a threat to human health, food security, and the future of the cocoa industry. Therefore, advances in agronomic and genetic methods that reduce this heavy metal in plants are necessary. In the present study, an *in silico* metadata analysis was carried out to identify genes involved in cadmium absorption to propose them as a target for genetic improvement through modern precision biotechnology techniques as genetic editing. We found 26 proteins, 88% of them reported in *Arabidopsis thaliana*, and 38% in *Oryza sativa*, which belong to 8 protein families. The remaining members of each family were included for analysis. Thus, expression data obtained from the Genevestigator software were evaluated for 44 genes of *A. thaliana* and 37 of *O. sativa*, and it was found that the *TcNRAMP5* gene of *T. cacao*, "orthologs" of *OsNRAMP5* of *O. sativa*, is closely related to the root absorption process of Cd. Consequently, it was postulated as a gene target for genetic editing to decrease Cd concentration in cocoa plants. A genetic editing methodology based on CRISPR/Cas9 system was structured using *Agrobacterium tumefaciens* mediated genetic transformation and protoplast transfection. Furthermore, this research structured a methodology for the establishment and generation of *in vitro* plants through indirect somatic embryogenesis and protoplast isolation, as a fundamental basis of the proposed modern biotechnological platform, which leads to the improvement of genetic accuracy in cocoa.

Palabras clave: CRISPR/Cas9, edición genética, metales pesados, mejoramiento nutricional, biotecnología molecular, cadmio.

1. Introducción

Theobroma cacao es una planta diploide ($2n=20$) perteneciente a la familia Malvaceae, subfamilia Byttnerioideae, cultivada principalmente en regiones tropicales para la obtención de sus semillas como materia prima en la elaboración de chocolate y derivados [1-3]. El árbol de cacao es nativo de la Amazonía superior, incluyendo Brasil, Colombia y Ecuador (ríos Napo, Putumayo y Amazonas) [4-6]. Aunque establecer el origen de su proceso de domesticación ha sido controversial e inicialmente se creía que este se habría llevado a cabo por los Mayas en Mesoamérica, hace 1500 años [6,7], en la actualidad se reconoce que dicho proceso se dio en el Holoceno medio, entre 5450 y 5300 años atrás, en la Amazonia superior al norte de Suramérica [5].

Según la Organización Internacional del Cacao (ICCO), se estima que para el año cacaotero 2018/2019 se produjeron alrededor de 4.780.000 toneladas de grano de cacao a nivel mundial, siendo los países africanos los mayores productores (Costa de Marfil, Camerún, etc.) con el 76,2% de la producción, seguidos por los productores latinoamericanos (Ecuador, Brasil, Colombia, etc.) con el 17,5% y, el 6,3% restante se le atribuye a países de Asia y Oceanía (Indonesia, Papúa Nueva Guinea, etc.) [2,8]. El mercado del grano de cacao para el 2018 reportó ingresos que ascienden a los US \$9.940 millones de dólares, y se espera que crezca a una tasa anual del 7,3% entre el 2019 y el 2025 [9]. Su uso en la producción de chocolate, manteca de cacao, cosméticos y otros productos impulsan un mercado de aproximadamente US \$44.500 millones de dólares [10], posicionando al cacao como cultivo potencial de explotación a gran escala. En Colombia se produjeron 59,790 toneladas de grano de cacao en el año 2019, siendo Santander el departamento que registró mayor producción, 25,158 toneladas, correspondiente al 42% de la producción nacional del año en cuestión, seguido por Antioquia y Arauca que registraron 5,259 (8,8%) y 4,546 (7,6%) toneladas de grano de cacao respectivamente [11]. En la actualidad y dentro del marco del post-conflicto, se ha postulado al cacao como un cultivo clave para programas sociales relacionados con la lucha contra la pobreza rural, paz en regiones de post-conflicto y sustitución de cultivos ilícitos [12]. No obstante, dentro de las limitaciones a las que se ha enfrentado la industria del cacao en las últimas décadas, que incluyen factores bióticos y abióticos, la carga de cadmio en sus granos y derivados representa uno de los principales retos actuales, amenazando la seguridad alimentaria, la salud humana y el futuro de la industria [13].

El cadmio es un metal pesado no esencial que se encuentra de forma natural en suelos, proveniente de la actividad volcánica [13]. Adicionalmente, el uso de combustibles fósiles, el vertimiento de desechos por parte de las industrias metalúrgicas y las malas prácticas agrícolas han propiciado su propagación en el medio ambiente [14], favoreciendo una mayor acumulación en los suelos y generando dos líneas de impacto: i) interfiriendo en el desarrollo de las plantas y, por ende, en el rendimiento de los cultivos de interés agrícola; ii) causando efectos adversos a la salud de humanos y animales, ya que una vez es absorbido y acumulado en tejidos vegetales ingresa a la cadena alimenticia cuando es consumido [15]. El envenenamiento por acumulación de cadmio puede llegar a tener graves consecuencias en la salud humana, siendo el causante de patologías relacionadas a la osteoporosis, enfermedades renales, enfermedades neurodegenerativas (p. ej. Parkinson y Esclerosis Lateral Amiotrófica (ELA)), enfermedades cardiovasculares y cáncer. [14, 16-18].

Para la población no fumadora el consumo de alimentos representa el 90% de la exposición a Cd, y de dicho consumo, los granos de cacao y el chocolate pueden aportar del 4 al 26% y del 15 al 92% respectivamente. Aunque no hay a la fecha datos sobre la concentración de Cd en productos a base de cacao capaz de afectar a la salud humana después de ingerirlos, el comité mixto FAO/OMS de expertos en aditivos alimentarios estableció una ingesta mensual tolerable de 0,025 mg/kg de peso corporal, mientras que el panel sobre contaminantes en la cadena alimentaria de la Autoridad Europea de Seguridad Alimentaria (EFSA) propuso una ingesta mensual tolerable de 0,01 mg/kg de peso corporal para garantizar la protección de todos los consumidores [19]. Aunque en Colombia la ingesta de cacao es baja, en países de la Unión Europea (UE) como Alemania, Países Bajos, Suiza, etc., se consume cerca del 50% de la producción mundial de cacao [9,20,21]. Por ello, la Comisión de Regulación de la UE bajo la regulación N° 488 del 2014, estableció los valores máximos de concentración de cadmio permitidos en productos

comestibles, entre ellos los productos a base de cacao. Dichos niveles oscilan entre 0,10 y 0,8 miligramos de Cd por kilogramo de producto. Esta regulación entró en vigencia el 1 de enero de 2019, por lo que desde esa fecha no se pueden comercializar productos que no cumplan con lo allí establecido [9,20]. Lo anterior pone en riesgo el mercado potencial para la industria del cacao colombiano que empieza a ser reconocido a nivel internacional, por su distinción de “Fino y de Aroma” otorgada por la ICCO [22,23]

Se ha encontrado que, para el grano de cacao producido en Latinoamérica y Centroamérica, los valores de concentración de Cd llegan hasta los 1,4 mg/kg [20]. En contraste, los valores arrojados para contenido de Cd en granos de cacao originarios de Asia, el oeste y este de África, fueron de 0,33, 0,09 y 0,5 mg/kg respectivamente, mostrando una gran diferencia con los valores arrojados para Latinoamérica y Centroamérica. En otro estudio desarrollado en Perú, se encontró que, algunas variedades de cacao podían llegar a tener concentraciones de Cd hasta de 12,56 mg/kg de cacao en polvo [13,20,24,25]. En Colombia se realizó una evaluación de la concentración de Cd en chocolate con un contenido superior al 30% de cacao, y arrojó un valor promedio de 4,0477 mg/kg [26]. En Ecuador se halló que la concentración promedio de Cd en grano fue de 0,90 mg/kg en 159 muestras de diferentes regiones, coincidiendo con valores similares reportados en otros países de Latinoamérica [27, 28]. Todo lo anterior posiciona al cacao producido en América como el de mayor contenido de Cd por kilogramo en el mundo [25]. Estos valores no solo exceden los límites máximos establecidos por la regulación de la UE, sino que ponen en riesgo la salud de los consumidores locales y el crecimiento del mercado internacional del cacao.

En atención a todo lo anterior, disminuir la concentración de Cd en plantas y derivados del cacao es absolutamente necesario. Dentro de las estrategias más prometedoras se encuentra la disminución de la absorción por los árboles, estrategia que puede desarrollarse desde diferentes metodologías: i) Manejo de fertilizantes, calidad del agua y ciclos de inundación-sequia ii) Adición de elementos que modifiquen condiciones fisicoquímicas del suelo que afecten la biodisponibilidad de Cd y iii) Adición de microorganismos u otras especies de plantas capaces de secuestrar Cd [28]. Además, diversos estudios han demostrado que la acumulación y tolerancia a metales pesados están asociadas a factores genéticos heredables [29,30], por lo que comprender las bases genéticas asociadas al proceso de acumulación y particularmente de absorción de Cd por la membrana plasmática en células de raíz, representa otra estrategia prometedora para la disminución de la absorción empleando técnicas de ingeniería genética. En el presente estudio se realizó un análisis *in silico* de metadatos para identificar genes clave en el proceso de absorción de Cd en *T. cacao*, y proponer una estrategia biotecnológica que permita disminuir las concentraciones de Cd en granos de cacao.

2. Materiales y métodos

2.1. Obtención de Metadatos

Se ejecutó una búsqueda de información literaria de forma sistemática en 4 motores de búsqueda o bases de datos seleccionadas previamente (Google Scholar, SCOPUS, SpringerLink y EBSCOhost). El procedimiento se realizó en aras de conseguir la mayor información, de forma individual, acerca de la distribución en Colombia y el mundo de *T. cacao*, su genoma, genes reportados en plantas modelo o de interés comercial relacionados con la absorción, transporte y almacenamiento de cadmio, investigaciones sobre cultivo *in vitro* e ingeniería genética desarrollados en cacao, etc. Se definieron términos de búsqueda específicos para cada tema, los cuales se combinaron con operadores booleanos (AND, OR, según permitiera el buscador) y un periodo de tiempo (1990 a 2020) para hacer más específica y profunda la búsqueda (**Tabla S1**). Las referencias arrojadas por cada búsqueda fueron filtradas, revisadas y seleccionadas según su contenido y lo que a nuestro criterio fuese información útil. Finalmente se dio paso a la construcción de fichas bibliográficas con la información relevante encontrada en cada una de las referencias, lo cual permitiría tener la información disponible de forma más rápida y práctica.

2.1.2. Identificación de genes-proteínas involucradas en la absorción, transporte y almacenamiento de cadmio (Cd) en plantas

Con los resultados obtenidos de los metadatos se construyó un listado de genes-proteínas que participan en la absorción, transporte o almacenamiento de Cd en plantas. La lista fue nutrida con datos funcionales encontrados en los trabajos de investigación y *UniProt Knowledgebase* (UniProtKB) (<https://www.uniprot.org/>) (**Tabla 1**). Posteriormente, las proteínas encontradas fueron agrupadas de acuerdo a la familia proteica a la que pertenecen. Algunos miembros de cada familia fueron adicionados después de hacer una revisión de cada una de ellas. Todas las proteínas fueron revisadas en el *NCBI* (<https://www.ncbi.nlm.nih.gov/>) para la obtención de sus números de acceso, tanto para *Arabidopsis thaliana* como para *Oryza sativa*, ya que estas especies al ser plantas modelo, de interés comercial y que para el caso del arroz que acumula cadmio, tendrían mayor disponibilidad de datos e información en las bases de datos. Finalmente, se usó *The Bologna Annotation Resource* (<https://bar.biocomp.unibo.it/bar3/>) y *UniprotKB* para identificar las anotaciones ontológicas (GO) relacionadas con la ubicación subcelular de cada proteína.

2.2. Construcción de las redes de Interacción Proteína-Proteína (IPP) de *A. thaliana* y *O. sativa* para Familias de Proteínas Asociadas a Cadmio (FPAC)

Las redes IPP fueron creadas sometiendo los números de acceso de las proteínas identificadas al software *STRING* version 11.0 (<https://string-db.org/>), una base de datos de interacciones proteicas conocidas y predichas. Las asociaciones directas (físicas) e indirectas (funcionales) se obtienen de siete fuentes: minería de textos, contexto genómico, experimentos de alto rendimiento (conservados), co-expresión, co-ocurrencia, bases de datos y *gene fusion*.

2.3. Análisis de expresión diferencial de FPAC en *A. thaliana* y *O. sativa*

Los datos de expresión para los genes de las FPAC de *A. thaliana* y *O. sativa* fueron obtenidos del software *Genevestigator*® (<https://genevestigator.com/gv/>), usando los nombres de las proteínas, GeneID's, números de acceso, entre otros identificadores. *Genevestigator*® contiene datos transcriptómicos cuidadosamente seleccionados, curados, anotados y normalizados de repositorios públicos. Así, los datos de *Affymetrix Arabidopsis ATH1 genome array* y *Affymetrix Rice Genome Array* fueron filtrados individualmente, teniendo en cuenta genotipo (*Col-0* para *A. thaliana* e *IR64* y *Nipponbare* para *O. sativa*), órgano, estímulo y característica global de la muestra.

2.3.1. Análisis de expresión diferencial de las FPAC de *A. thaliana* y *O. sativa* en brote, raíces y hoja no tratadas

Se extrajeron datos de expresión de las FPAC usando la herramienta *Anatomy* de *Genevestigator*® para muestras *wild-type*, no tratadas y no infectadas (estímulo), tanto de *A. thaliana* como de *O. sativa*, esto se hizo para raíces (*Roots*), brote (*Shoot*) y hoja (*Leaf*). La herramienta *Anatomy* arroja valores de intensidad de expresión (\log_2) de los genes de interés en diferentes categorías anatómicas incluyendo órganos, tejidos y cultivos celulares.

2.3.2. Análisis estadístico

Los datos obtenidos para cada uno de los órganos, de manera independiente y para cada una de las familias, fueron evaluados mediante un Análisis de Varianza (ANOVA) de un solo factor y una Prueba F de Fisher, para así identificar si existían o no diferencias significativas entre las medias de expresión intrafamilias o posteriormente, interfamilias. En el software *Minitab 18* (<https://www.minitab.com/>) se ejecutó el ANOVA con un nivel de significancia $\alpha = 0,05$ y un nivel de confianza del 95%. Adicionalmente se generaron gráficos de barra para visualizar las diferencias de expresión encontradas entre miembros de una misma familia a través del software *GraphPad Prism 8* (<https://www.graphpad.com/>).

2.3.3. Análisis de expresión diferencial de las FPAC de *A. thaliana* y *O. sativa* en raíces tratadas con metales

Las proteínas expresadas principalmente en raíz fueron seleccionadas y se extraeron sus datos de expresión para muestras *wild-type* tratadas con Cd y Fe para *A. thaliana*, Cd y Pb para *O. sativa* (*IR64*) en la herramienta

Perturbations de Genevestigator®. La herramienta *Perturbations* proporciona un resumen de respuestas de expresión génica a una amplia variedad de perturbaciones, y tiene como propósito identificar condiciones que pueden causar una regulación hacia arriba o hacia abajo de los genes de interés, esto lo hace comparando valores de expresión de muestras experimentales (tratadas) versus muestras control. El resultado refleja la magnitud de la diferencia entre el promedio de la expresión en las muestras experimentales y el promedio en las muestras de control. Los valores obtenidos fueron posteriormente visualizados en la herramienta *Expression* del software *Heatmapper* (<http://www.heatmapper.ca/>).

2.4. Análisis de homologías

Las proteínas principalmente expresadas en raíz y que mostraron una respuesta significativa a los tratamientos con metales, incluyendo Cd, fueron sometidas a un BLASTp en las plataformas *BLAST®* del NCBI y *Cocoa Genome Hub* (<https://cocoa-genome-hub.southgreen.fr/>) en búsqueda de proteínas con secuencias homólogas en *T. cacao*, otras plantas de interés comercial (*Zea mays* y *Arachis hypogaea*) e hiperacumuladoras de cadmio (*Arabidopsis halleri* y *Sedum alfredii*). Aquellas proteínas homólogas en *T. cacao* y compartidas en los resultados de BLASTp para *A. thaliana* y *O. sativa*, fueron sometidas en conjunto con los homólogos resultantes de las otras plantas, a un alineamiento de secuencias proteicas por medio del algoritmo *ClustalW* en el software *MEGAX* versión 10.1.8 (<https://www.megasoftware.net/>), teniendo en cuenta que las proteínas pertenecieran a la misma familia proteica. Con los alineamientos obtenidos se generaron árboles filogenéticos por el método de *Maximum Likelihood* basado en modelo de matriz JTT y con valores de *Bootstrap* obtenidos de 1500 réplicas.

2.4.1 Estructura genética y arquitectura de Motivos conservados

Se identificó la estructura de los genes y se predijo la arquitectura de Motivos conservados de las proteínas que hacen parte del análisis filogenético usando los software en línea *Gene Structure Display Server* (GSDS2.0 <http://gsds.cbi.pku.edu.cn/>) y *MEME* versión 5.1.1, respectivamente (<http://meme-suite.org/tools/meme>).

2.5. Análisis de metodologías para el establecimiento y regeneración in vitro de plantas de cacao

Se revisaron y filtraron las referencias relacionadas con metodologías de regeneración y establecimiento *in vitro* en *T. cacao* obtenidas de la búsqueda bibliográfica. Para proponer la metodología a ejecutar se tuvieron en cuenta criterios como tipo de explante, medios de cultivo, porcentaje de éxito en generación de callos, generación de embriones, tiempos de cultivo, subcultivo y endurecimiento.

2.6. Estrategia de Edición genética

2.6.1. Diseño de ARN guía (sgRNA)

Con base en los resultados obtenidos a partir de los análisis de expresión diferencial, la búsqueda de proteínas homólogas en *T. cacao*, sus relaciones filogenéticas, estructuras genéticas, arquitectura de Motivos conservados y la información bibliográfica obtenida para cada proteína, se seleccionó un (1) gen objeto de edición. El diseño del *sgRNA* se realizó usando la herramienta *CRISPR Site Finder* del software *Geneious Prime* versión 2020.2 (<https://www.geneious.com/>). El gen de interés y el genoma de *T. cacao*, variedad criollo (v2), fueron descargados del NCBI para el reconocimiento de la secuencia PAM (NGG) y sitios *off-target* respectivamente. El *sgRNA* se seleccionó de acuerdo a los siguientes criterios: menor cantidad de sitios *off-target*, ubicación dentro de una región codificante del gen, mejor coeficiente de actividad, mayor porcentaje de especificidad y al menos un sitio de corte para enzimas de restricción.

2.6.2. Elección del vector

El vector fue elegido según la guía de MilliporeSigma “*Genome Editing in Plants with CRISPR/Cas9*” (<https://www.sigmaaldrich.com>) de acuerdo a la especie, el tipo de plásmido y el gen marcador de selección.

2.6.3. Metodología de edición

Se estructuró una propuesta metodológica para desarrollar en una siguiente fase experimental (FASE 2), la edición genética de plantas de cacao. Esta propuesta se elaboró con base en los resultados obtenidos de la búsqueda bibliográfica y el conocimiento propio de los autores involucrados en el trabajo. Se plantean 2 alternativas para realizar el mejoramiento genético de la planta vía edición genética: por medio de la transformación mediada por *Agrobacterium tumefaciens* de embriones primarios de cacao provenientes de explantes florales y por transfección de protoplastos foliares con la maquinaria del sistema de edición *CRISPR/Cas9*.

3. Resultados

3.1. Obtención de Metadatos

De la búsqueda literaria se extrajeron 276 trabajos de investigación distribuidos entre los criterios de búsqueda utilizados para los diferentes ítems consultados (**Tabla S1**). Los trabajos fueron revisados y organizados según su contenido e información y su relación o no con los ítems especificados para la búsqueda. De cada trabajo se extrajo la información que se consideraba importante y relevante, esta se consignó en fichas incluyendo el título, autores, objetivo, resultados, etc., para luego llegar a evaluar su pertinencia y posible uso como fuente bibliográfica en el presente trabajo. Finalmente se dispuso de aquellos documentos con información útil y que se ajustara a los resultados que se esperaban obtener para la búsqueda, estos fueron finalmente utilizados como fuentes bibliográficas.

3.1.2. Proteínas relacionados con absorción, transporte y almacenamiento de Cadmio

Los trabajos de investigación extraídos de las bases de datos previamente seleccionadas, fueron cuidadosamente revisados con el fin de identificar genes-proteínas relacionadas con absorción, transporte y almacenamiento de Cd en plantas. Los datos obtenidos fueron consignados y su función fue complementada con información obtenida de *UniProtKB* (**Tabla 1**).

Se obtuvo un listado de 26 proteínas para 12 especies de plantas incluyendo plantas modelo, de interés comercial e hiperacumuladoras de Cd. Aproximadamente el 88% (23) y el 38% (10) de los genes-proteínas hallados se registraron para *A. thaliana* y *O. sativa* respectivamente, por lo que estas dos especies fueron usadas para los análisis posteriores. A todas las proteínas se les atribuyó una función relacionada con el metabolismo del cadmio, principalmente asociadas a membrana plasmática y/o vacuola. Un total de 24 proteínas, de las 26 registradas, fueron clasificadas en 8 familias proteicas existentes: ABCC (ATP-Binding Cassette), NRAMP (Natural Resistance-Associated Macrophage Protein), ZIP (Zrt-and Irt-regulated transporter), CDF (Cation Diffusion Facilitator), CAX (Cation-Proton Exchanger), COPT (Copper Transporter), OPT (Oligopeptide Transporter) y HMA (Heavy Metal ATPase). Cada una de las familias mencionadas están asociadas al transporte de iones esenciales como Ca, Zn, Co, Mn y Fe. Sin embargo, su baja selectividad contribuye a que iones no esenciales, como los de Cd, ingresen a las células vía simplástica, y posteriormente sean acumulados en vacuolas de diferentes tejidos, exponiendo así, tanto a la planta como a quien la consume (o sus derivados) a los posibles efectos adversos mencionados anteriormente [31, 32]. Particularmente el miembro IRT1 de la familia ZIP, un transportador de hierro capaz de transportar otros iones metálicos, ha sido de los más reportados de acuerdo a los trabajos encontrados, seguido por miembros de la familia NRAMP y HMA, tales como NRAMP5 y HMA3 respectivamente, sugiriendo una estrecha relación de estos con el metabolismo de Cd en plantas.

Proteína	Especie	Función	Referencias
IRT1	<i>A. thaliana</i> <i>Arachis hypogaea</i> <i>Nicotiana tabacum</i>	Transportador de membrana plasmática para hierro (Fe), zinc (Zn), cobalto (Co), manganeso (Mn) y cadmio (Cd) en la capa epidérmica de la raíz.	[33-42]
ZNT1 ZIP4*	<i>Noccaea caerulea</i> <i>A. thaliana</i>	Transportador de membrana plasmática para zinc y cadmio.	[29, 43]

NRAMP3 NRAMP4	<i>A. thaliana</i> <i>N. caerulea</i>	Transportador vacuolar de hierro, manganeso y cadmio (funcionan redundantemente).	[40,44-47]
ZAT-1	<i>A. thaliana</i>	Transportador de zinc y cadmio; relacionado con el secuestro vacuolar de estos metales.	[29, 48]
OPT3	<i>A. thaliana</i> <i>A. hypogaea</i>	Transportador de membrana con capacidad para cargar hierro y cadmio al floema (facilitando la recirculación del xilema al floema).	[44]
NRAMP5	<i>A. hypogaea</i> <i>O. sativa</i> <i>Zea mays</i> <i>Brassica juncea</i> <i>Malus baccata</i>	Transportador de membrana plasmática para hierro y cadmio.	[41,44,49-52]
HMA4 HMA2	<i>Arabidopsis halleri</i> <i>O. sativa</i> <i>A. thaliana</i>	Transportador de zinc y cadmio. Se encarga de la traslocación de raíz a tallo y carga al xilema.	[38, 53-57]
HMA3	<i>O. sativa</i> <i>A. thaliana</i> <i>A. halleri</i> <i>Sedum plumbizincicola</i>	Transportador vacuolar de zinc y cadmio. Se ubica principalmente en raíz, evitando el paso de raíz a brotes y granos.	[38, 49,54-56,58,59]
PDR8	<i>A. thaliana</i>	Confiere resistencia al cadmio y al plomo, probablemente como una bomba de eflujo de Cd o conjugados de Cd (principalmente vello radicular y epidermis).	[60-62]
ABCC3 ABCC1 ABCC2	<i>A. thaliana</i>	Bomba para conjugados de glutatión S. Permite el transporte de cadmio conjugado.	[41,55,63,64]
DTX1	<i>A. thaliana</i>	Involucrado con la detoxificación de cadmio, actuando como bomba de eflujo, extruyendo del citoplasma al apoplasto.	[65]
MTP1	<i>A. thaliana</i>	Media la acumulación de zinc y cadmio en vacuolas de raíces.	[65]
ZNT5	<i>N. caerulea</i>	Puede estar relacionado con la acumulación de zinc y cadmio en células de la epidermis y células de la envoltura del haz (aún no se esclarece).	[65]
ZIP2 ZIP3	<i>A. thaliana</i> <i>O. sativa</i>	Media la absorción de zinc, cobre y cadmio.	[38]
CAX2 CAX4	<i>A. thaliana</i> <i>Brassica rapa</i> <i>N. tabacum</i>	Intercambiador catión/protón vacuolar. Trasloca Ca^{2+} y otros iones metálicos utilizando el gradiente de protones.	[38,66,67]
MRP7	<i>O. sativa</i> <i>B. rapa</i>	Transportador vacuolar ABC para S-conjugados de glutatión.	[38]
MTP3	<i>O. sativa</i> <i>B. rapa</i>	Transportador de vacuolar de metales, incluyendo cadmio.	[38]
COPT5	<i>A. thaliana</i> <i>O. sativa</i> <i>Nicotiana attenuata</i>	Involucrado en el transporte vacuolar de cobre (Cu) y cadmio.	[38]
PCR1	<i>A. thaliana</i>	Transportador de la membrana plasmática. Participa en la resistencia al cadmio independiente de glutatión.	[68]
NRAMP1	<i>A. thaliana</i> <i>O. sativa</i>	Media la acumulación de Cd en raíces.	[40,69]

Tabla 1. Proteínas relacionadas con el metabolismo de cadmio. * Genes ortólogos. La información sobre funciones aquí contenida fue confirmada con las anotaciones de UniProtKB (<https://www.uniprot.org/>).

Los miembros restantes de las familias mencionadas arriba, para *A. thaliana* y *O. sativa*, fueron incorporados a los análisis [68,70-78]. Finalmente, usando *BAR 3.0* y *UniProtKB* se accedieron a los datos de ubicación subcelular para cada miembro de las FPAC (**Datos S1** y **Datos S2**). Aquellas familias de proteínas cuyos miembros estuviesen

principalmente ubicados en vacuola u otras locaciones, fueron descartadas (CAX y ABCC), y las familias restantes (ABC, ZIP, COPT, OPT, NRAMP y HMA) localizadas principalmente en membrana plásmatica, fueron seleccionadas para los análisis posteriores.

3.2. Construcción de las redes de Interacción Proteína-Proteína (IPP) de *A. thaliana* y *O. sativa* para FPAC

Las redes de Interacción Proteína-Proteína construidas tanto para *A. thaliana* como *O. sativa* confirmaron, con un $p\text{-value} < 1 \times 10^{-16}$, que todas las proteínas seleccionadas están al menos parcialmente conectadas como un grupo, es decir, se relacionan dentro de una posible red metabólica asociada al metabolismo de iones metálicos (**Figura S1** y **Figura S2**). Para *A. thaliana* (**Figura S1**) se obtuvieron 210 interacciones derivadas de 44 nodos (proteínas) y un promedio general de 10 interacciones por nodo. El nodo con más interacciones fue el correspondiente a la proteína AtMTPA2 con 67 interacciones con 27 proteínas. Por otro lado, para *O. sativa* (**Figura S2**) la red IPP mostró 148 interacciones entre las 37 proteínas que la conforman. El número promedio de interacciones por proteína fue de 8, y la proteína con más interacciones dentro de la red fue la OsMTP1 mostrando 81 interacciones con 26 proteínas diferentes. Finalmente, el coeficiente de agrupación local promedio para las redes fue cercano a 0,5 (0,492 y 0,479 para *Arabidopsis* y arroz respectivamente) lo cual sugiere que, en promedio, el 50% de los nodos de la red están conectados entre sí.

3.3. Análisis de expresión diferencial de las FPAC de *A. thaliana* y *O. sativa*

3.3.1. Análisis de expresión diferencial de las FPAC de *A. thaliana* y *O. sativa* en brotes, raíces y hojas no tratadas

Para identificar proteínas expresadas principalmente en raíz, se extrajeron los datos de expresión de la herramienta *Anatomy* de *Genevestigator*® y fueron sometidos a un ANOVA y una prueba F de Fisher. Los resultados del ANOVA para datos de raíz y cada familia por separado, permitió establecer la existencia de al menos una diferencia significativa en los promedios de expresión entre miembros de una misma familia. Por otra parte, el método de la Diferencia Menos Significativa (LSD) de Fisher permitió visualizar las diferencias significativas entre cada uno de los miembros. Así, aquellas cuyo promedio de expresión fue $\geq 80\%$ con respecto a la de mayor promedio, fueron seleccionadas para un análisis entre tejidos, comparando los valores promedio de expresión en raíz de dichas proteínas, con respecto a su expresión promedio en brotes y hojas.

Para *A. thaliana* las proteínas AtNRAMP1, AtMTPA2, AtMTP10, AtHMA4, AtHMA5, AtHMA7, AtIRT1, AtZIP2 y AtZIP3 exhibieron los mayores valores medios de expresión en raíz y exhibieron al menos un 10% de diferencia con respecto a los valores medios de expresión para brotes y hoja (**Figura 1**). AtNRAMP1 mostró mayores niveles de expresión dentro del grupo de proteínas mencionadas y, según los resultados obtenidos del ANOVA y por el método de la Diferencia Menos Significativa (LSD) de Fisher, AtNRAMP1 fue agrupada de forma independiente a las demás proteínas, mientras que AtHMA7 y AtZIP2 formaron otro grupo y mostraron diferencias significativas con respecto a AtHMA4, AtNRAMP4, ZIP3 y AtMTP10. Aunque no pudo compararse estadísticamente los resultados entre tejidos debido a las diferencias en los tamaños muestrales, las OPT mostraron una preferencia notable de expresión en brotes y hojas en comparación con raíz, mientras que las COPT no mostraron casi diferencias entre tejidos (**Datos S3** y **Datos S5**).

Para *O. sativa* (*Nipponbare*), las proteínas OsNRAMP5, OsNRAMP6, OsMTP7, OsHMA3, OsHMA5, OsCOPT1, OsZIP1, OsZIP5, OsZIP7 y OsZIP9 mostraron los mayores valores medios de expresión en raíz y con al menos un 10% de diferencia con respecto a los valores medios de expresión para brotes y hoja (**Figura 2**). La OsCOPT1, la OsZIP1 y OsNRAMP5 mostraron mayores niveles de expresión en raíz comparadas con las demás proteínas. Sin embargo, según los resultados obtenidos del ANOVA y por el método de la Diferencia Menos Significativa (LSD) de Fisher, la OsCOPT1 fue agrupada de forma independiente a las demás proteínas. Por otro lado, OsZIP1 y OsNRAMP5 también se ubicaron en un grupo totalmente independiente a OsNRAMP6, OsMTP7, OsHMA3, OsHMA5, OsZIP5, OsZIP7 y OsZIP9 (**Figura S3**).

Según los resultados obtenidos para *O. sativa* (IR64), las proteínas OsNRAMP5*, OsMTP7*, OsHMA4*, OsHMA5*, OsCOPT1* y OsZIP1* mostraron los mayores niveles expresión en raíz y al menos un 10% de diferencia con respecto a los valores medios de expresión para brotes y hoja (**Figura 3**). La OsCOPT1* y la OsZIP1* fueron las que se expresaron más en raíz comparadas con las demás proteínas. Pero, según los resultados arrojados por el ANOVA y el método de la Diferencia Menos Significativa (LSD) de Fisher, OsCOPT1* y OsZIP1* fueron agrupadas por separado, y de forma independiente de las demás proteínas. En adición, OsHMA4 y OsNRAMP5 también se ubicaron en un grupo conjunto con OsHMA5*, manteniendo niveles de expresión altos y significativamente diferentes de OsMTP7 y OsHMA5 (**Figura S4**) (**Datos S4** y **Datos S5**).

Las diferencias en expresión fueron significativas cuando se compararon entre sí (intrafamilia e interfamilia) con una confianza del 95% ($\alpha = 0,05$) y un *p-value* igual a cero (0).

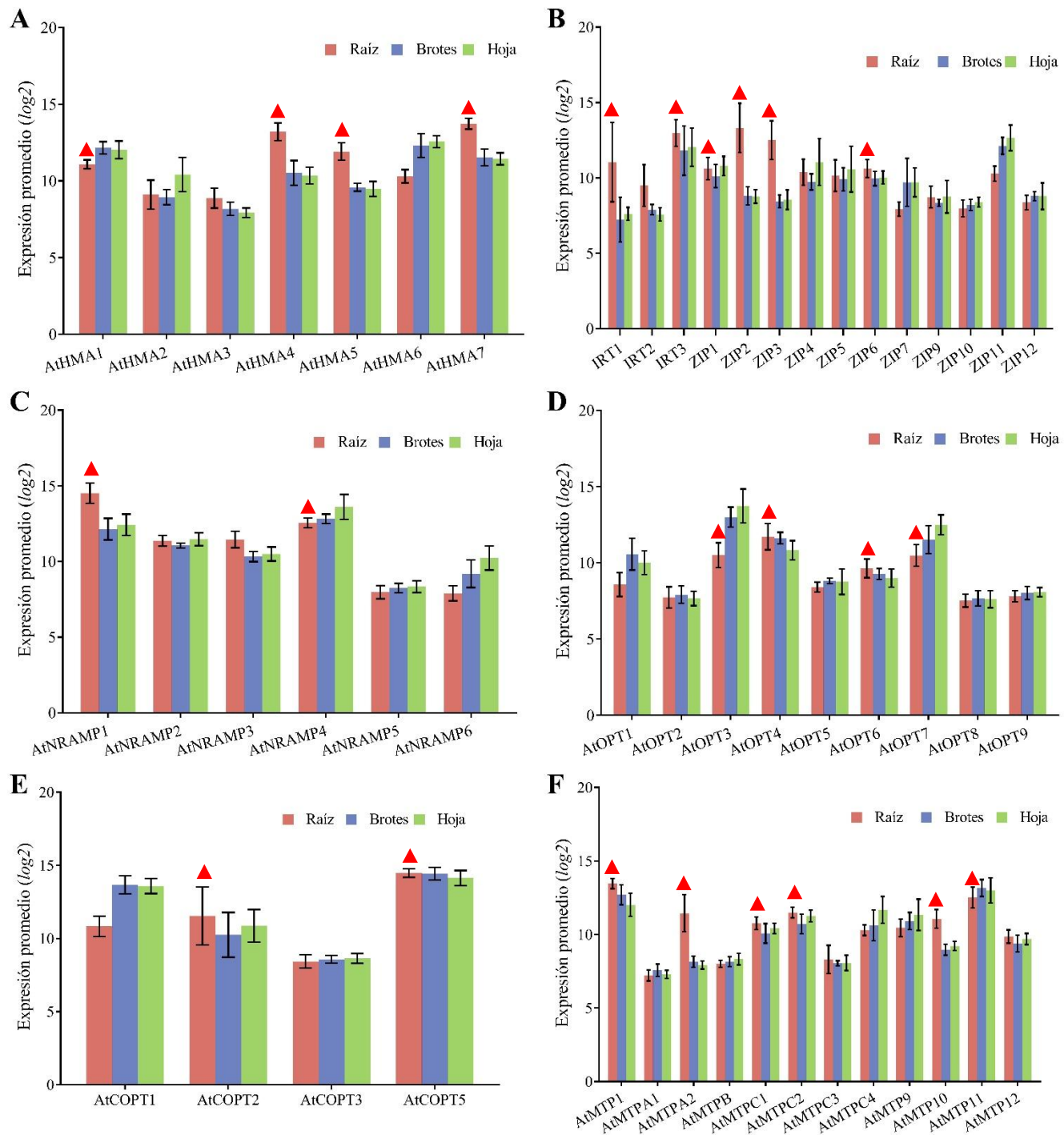


Figura 1. Niveles de expresión de las FPAC en raíces, brotes y hoja de plantas de *A. thaliana* no tratadas (*Col-0*). **(A)** Familia de proteínas HMA. **(B)** Familia de proteínas ZIP. **(C)** Familia de proteínas NRAMP. **(D)** Familia de proteínas OPT. **(E)** Familia de proteínas COPT. **(F)** Familia de proteínas CDF. Los datos corresponden a los valores promedio de expresión \pm DE de 63, 11 y 31 muestras para raíz, brotes y hoja respectivamente. El triángulo rojo (\blacktriangle) señala las proteínas que mostraron una expresión significativamente mayor en raíz para cada familia.

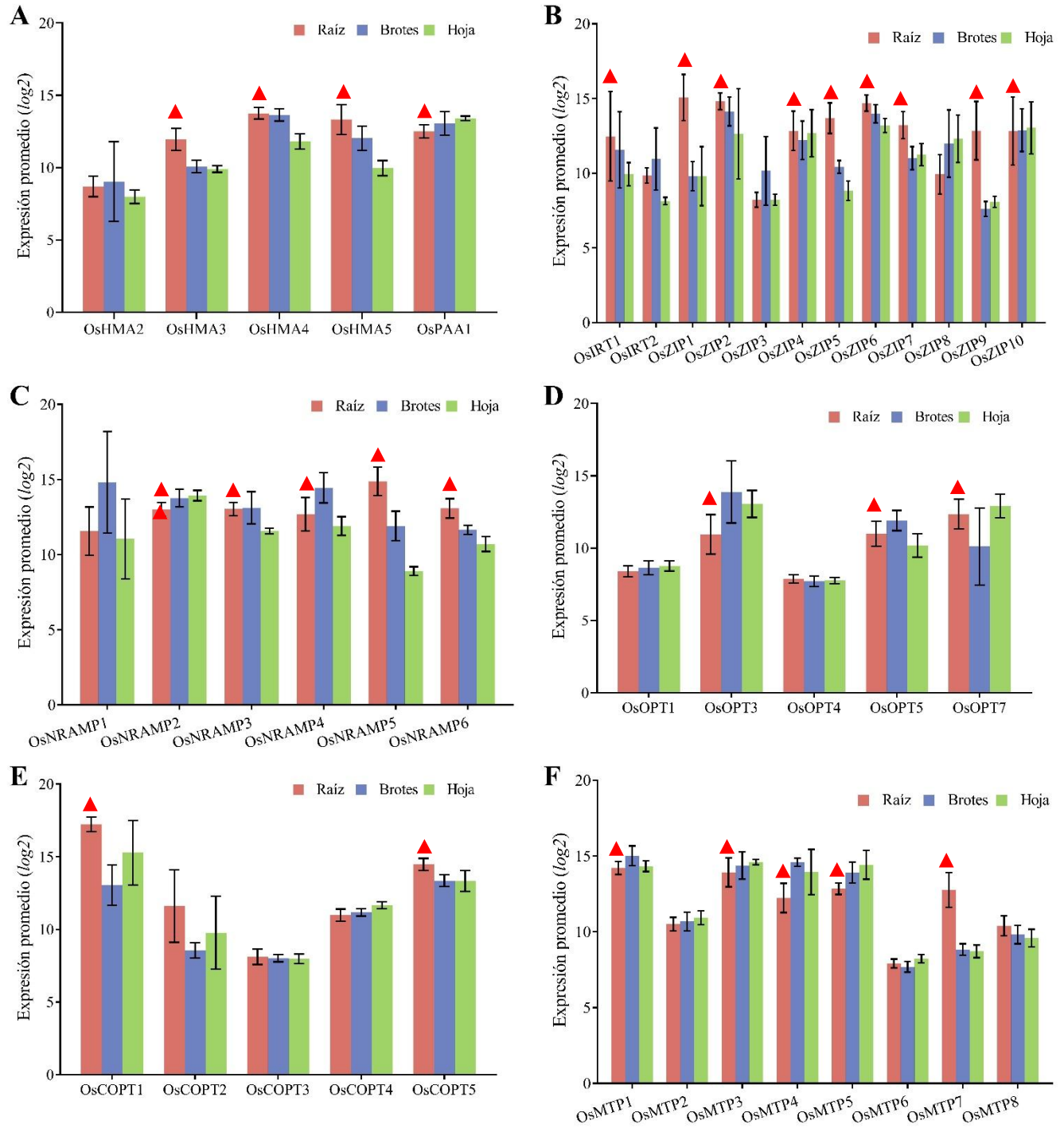


Figura 2. Niveles de expresión de las FPAC en raíz, brotes y hoja de plantas de *O. sativa* no tratadas (*Nipponbare*). **(A)** Familia de proteínas HMA. **(B)** Familia de proteínas ZIP. **(C)** Familia de proteínas NRAMP. **(D)** Familia de proteínas OPT. **(E)** Familia de proteínas COPT. **(F)** Familia de proteínas CDF. Los datos corresponden a los valores promedio de expresión \pm DE de 23, 15 y 9 muestras para raíz, brotes y hoja respectivamente. El triángulo rojo (\blacktriangle) señala las proteínas que mostraron una expresión significativamente mayor en raíz para cada familia.

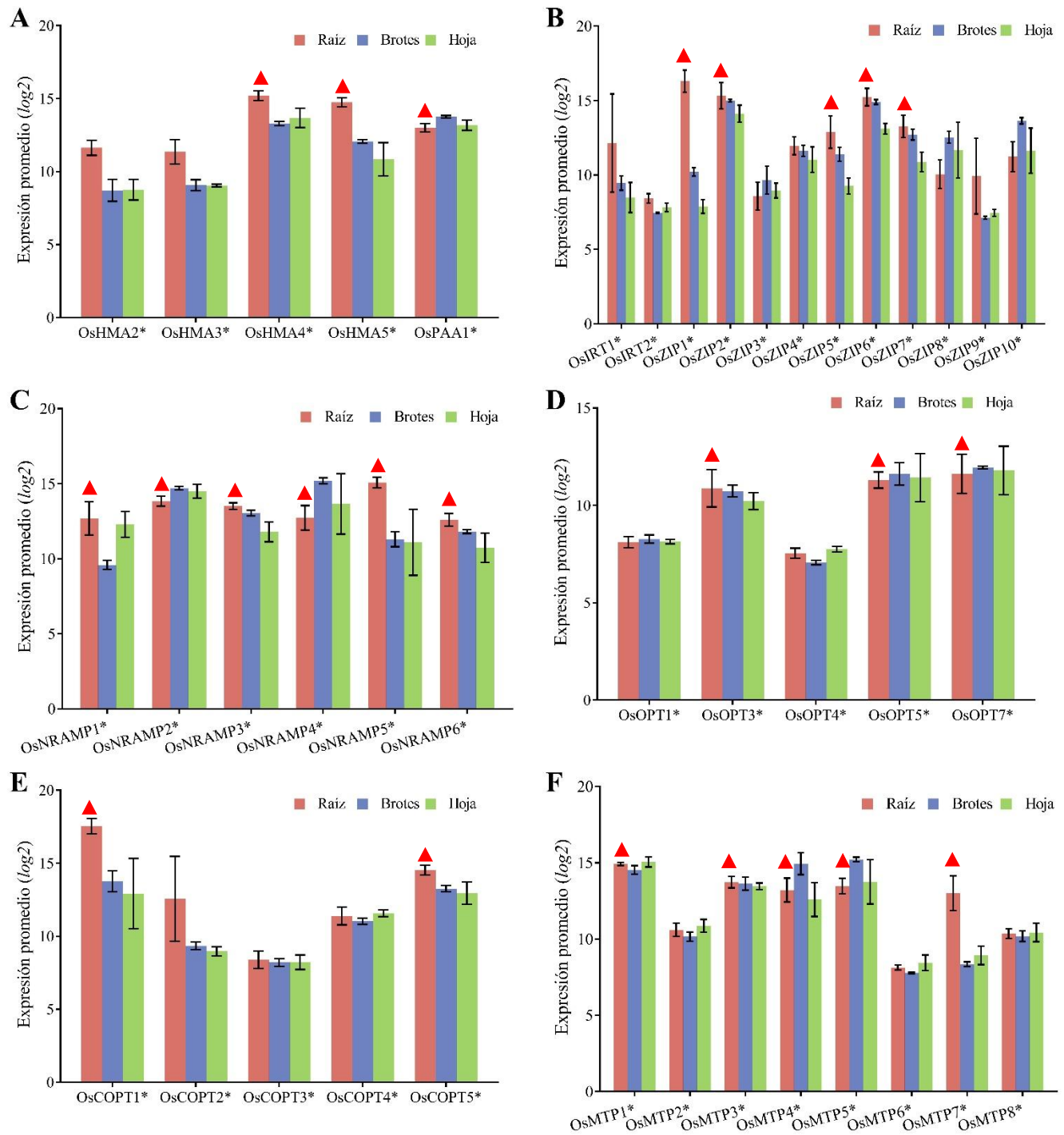


Figura 3. Niveles de expresión de las FPAC en raíz, brotes y hoja de plantas de *O. sativa* no tratadas (*IR64*). **(A)** Familia de proteínas HMA. **(B)** Familia de proteínas ZIP. **(C)** Familia de proteínas NRAMP. **(D)** Familia de proteínas OPT. **(E)** Familia de proteínas COPT. **(F)** Familia de proteínas CDF. Los datos corresponden a los valores promedio de expresión \pm DE de 9, 3 y 9 muestras para raíz, brotes y hoja respectivamente. El triángulo rojo (\blacktriangle) señala las proteínas que mostraron una expresión significativamente mayor en raíz para cada familia.

3.3.2. Análisis de expresión diferencial de las FPAC de *A. thaliana* y *O. sativa* en raíces tratadas con metales

Con el fin de identificar si las proteínas expresadas principalmente en raíz respondían a estímulos por exceso o deficiencia de metales, fueron extraídos datos de expresión bajo perturbaciones por exceso de cadmio (Cd), plomo (Pb) o deficiencia de hierro (Fe) de *Genevestigator*®. Para *A. thaliana* ninguna de las proteínas mostró regulación hacia arriba o hacia abajo cuando fueron sometidas a tratamientos con 200µM de ClCd_2 por 6h, teniendo en cuenta un $p\text{-value} \leq 0,05$ y un nivel de confianza del 95% ($\alpha=0,05$). No obstante, ligeras variaciones pudieron visualizarse en el mapa de calor (**Figura 4A**) para AtNRAMP1, AtMTPA2, AtHMA4 y AtIRT1. Para los tratamientos con deficiencia de hierro, el mayor tiempo de evaluación (Fe3/48-72h) arrojó diferencias significativas para AtNRAMP1, AtNRAMP4, AtHMA5, AtMTPA2, AtIRT1 y AtZIP3 con una clara regulación hacia arriba para todas, excepto AtZIP3, quien mostró regulación hacia abajo. El tratamiento con tiempo de evaluación medio (Fe2/12-24h) mostró diferencias significativas para AtNRAMP4 y AtZIP3, mientras que el tratamiento con menor tiempo de evaluación (Fe1/3-6h) no mostró diferencias significativas con respecto a los tratamientos control.

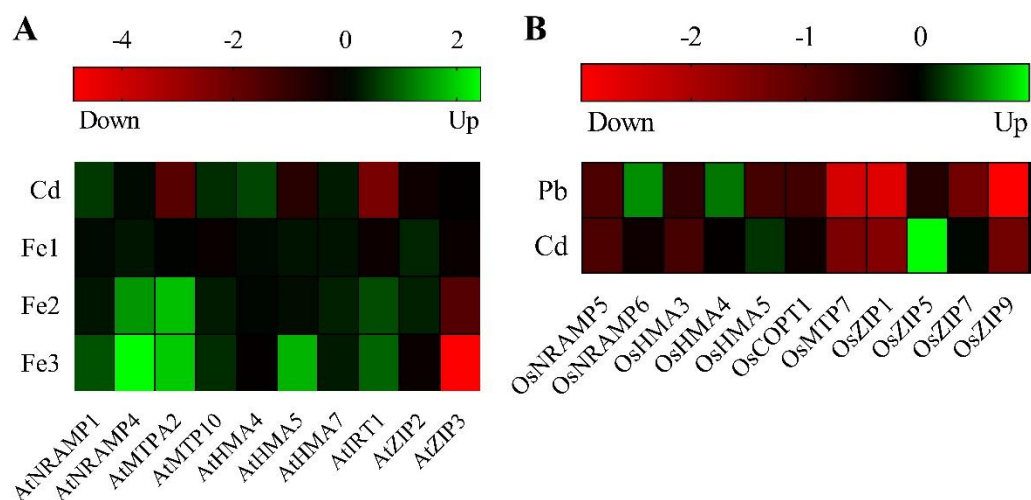


Figura 4. Mapa de calor de expresión en raíz de miembros de las FPAC bajo condiciones de deficiencia o exceso de metales. (A). Cambio en la expresión de miembros de las FPAC de *A. thaliana* expuestas a 200µM de ClCd_2 (Cd) por 6h, deficiencia de hierro (Fe) entre 3h a 6h (Fe1), 12h a 24h (Fe2) y 48h a 72h (Fe3), todos contra muestras control. (B). Cambio en las expresión de miembros de las FPAC de *O. sativa* (IR64) expuestas a 100µM de nitrato de plomo ($\text{Pb}(\text{NO}_3)_2$) y cloruro de cadmio (ClCd_2) durante 24h versus muestras control. Valores cercanos a cero (0) indican cambios leves en los niveles de expresión. Un color rojo indica regulación hacia abajo y el color verde regulación hacia arriba. Los datos obtenidos fueron soportados por valores de $p\text{-value}$ arrojados por la herramienta *Perturbations* de *Genevestigator*®.

Las muestras de raíz de *O. sativa* (IR64) que fueron sometidas a tratamientos con 100µM de cloruro de cadmio (ClCd_2) por 24h mostraron regulación, hacia arriba o hacia abajo, con valores $p\text{-value} \leq 0,05$ para OsNRAMP5, OsHMA3, OsZIP1 y OsZIP5, mientras que para el tratamiento con 100µM de nitrato de plomo ($\text{Pb}(\text{NO}_3)_2$) estas mismas proteínas aunque mostraron respuestas en su regulación, los valores de $p\text{-value}$ no soportaron dichos comportamientos, es decir, $p\text{-value} > 0,05$ (**Figura 4B**).

3.4. Análisis de homologías

Las proteínas que exhibieron una respuesta significativa a la deficiencia de Fe para *A. thaliana* y exceso de Cd para *O. sativa* fueron sometidas a un *BLASTp* con el fin de identificar homólogos en *T. cacao*. Se extrajeron aquellas cuyos valores E_value fueron más cercanos a 0, porcentajes de cobertura $> 70\%$ y porcentajes de identidad $> 45\%$. Tanto AtNRAMP1 como OsNRAMP5 arrojaron las mismas secuencias homólogas para *T. cacao*, correspondientes a TcNRAMP5 y TcNRAMP6 (E_value iguales a 0, porcentajes de cobertura $> 90\%$ y porcentajes de identidad $>$

50%). De la misma forma, AtZIP3 y OsZIP5 arrojaron las mismas secuencias homólogas en *T. cacao* (E-value cercanos a 0, porcentajes de cobertura $\geq 89\%$ y porcentajes de identidad $> 50\%$) correspondientes a TcZIP5 y TcZIP1, sugiriendo que estas proteínas presentan una estructura y función similar conservada en las tres especies de plantas. Adicionalmente el mismo análisis fue realizado en búsqueda de homólogos en otras plantas tanto de interés comercial (*Z. mays* y *A. hypogaea*) como hiperacumuladoras de Cd (*A. hallery* y *S. alfredii*). Los resultados arrojados por el *BLASTp* permitieron identificar homólogos de ambas familias proteicas en las 4 especies ya mencionadas. Para *Z. mays* se identificaron las ZmNRAMP3, ZmNRAMP5 y la ZmZIP1. En el caso de *A. hypogaea* resultaron la AhyNRAMP5, AhyNRAMP6 y AhyZIP5. AhNRAMP5, AhZIP3 y SaNRAMP5 para *A. hallery* y *S. alfredii* respectivamente.

3.4.1. Estructura genética y arquitectura de Motivos conservados

Con las secuencias de aminoácidos de los homólogos encontrados entre las 7 especies de plantas, se generó un árbol filogenético para cada una de las familias proteicas (ZIP y NRAMP), soportado por valores de *Bootstrap* obtenidos de 1500 réplicas, tal como se muestran en la parte izquierda de las **Figuras 5 y 6**. Adicionalmente la estructura genética y la arquitectura de los Motivos conservados soportaron las filogenias obtenidas, dando como resultado la agrupación de las proteínas en 3 distintos grupos para la familia NRAMP, mientras que para la familia ZIP las agrupaciones no fueron evidentes. Para las NRAMP se predijeron 16 Motivos conservados y, para las ZIP fueron 8 en total. En las NRAMP (**Figura 5**) el primer grupo (azul) está conformado por OsNRAMP5, ZmNRAMP5, TcNRAMP5 y AhyNRAMP5, las cuales tienen entre 11 y 13 exones en su estructura genética y, además, sus secuencias proteicas comparten los mismos 15 Motivos. Un segundo grupo (rojo) está conformado por AtNRAMP4 y AhNRAMP5, las cuáles poseen 3 y 4 exones respectivamente, y comparten entre sí 9 Motivos, únicamente difiriendo en el Motivo 15 que sólo está presente en AhNRAMP5. El tercer y último grupo de las NRAMP (verde) está formado por AtNRAMP1 y SaNRAMP5, cada una posee una estructura genética conformada por 11 exones y comparten 15 Motivos, ya que el Motivo 10 sólo está presente en los miembros del grupo 2. Se resalta la diferencia en el tamaño de la secuencia de los genes que conforman el grupo 1 en comparación con los del grupo 3, siendo estos últimos de menor tamaño de secuencia, pero teniendo el mismo número de exones y las proteínas compartiendo los mismos 15 Motivos.

Para la familia ZIP (**Figura 6**) el árbol generado fue pobremente soportado por bajos valores de *Bootstrap*, lo cual se vio reflejado tanto en la estructura genética de sus miembros como en la arquitectura de los Motivos conservados, debido a que no fue posible visualizar agrupaciones claras que compartiesen similitudes estructurales, a excepción de AtZIP3 y AhZIP3 que comparten una estructura genética similar (3 exones) y 6 de los 8 Motivos conservados predichos.

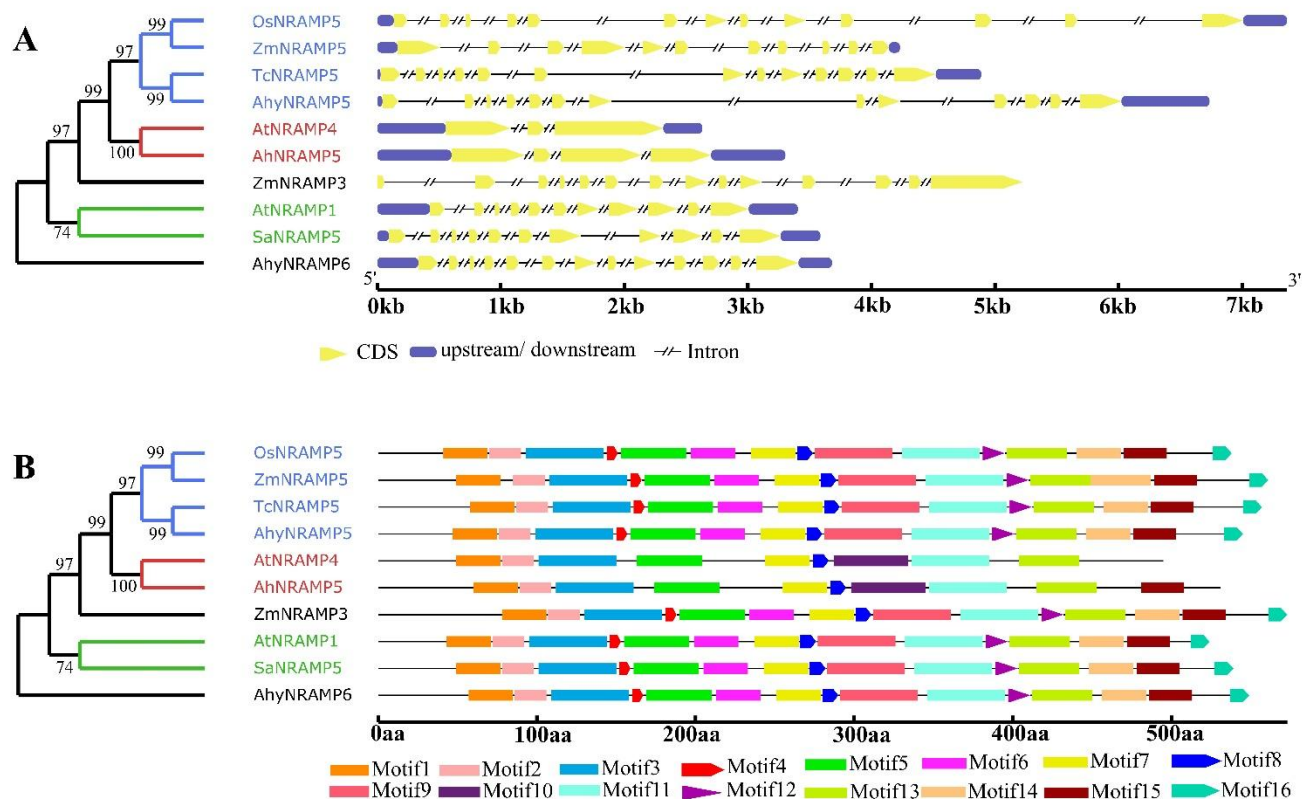


Figura 5. Estructura genética y arquitectura de Motivos conservados de miembros homólogos de la familia NRAMP de *T. cacao* (Tc), *O. sativa* (Os), *A. thaliana* (At), *Z. mays* (Zm), *A. hypogaea* (Ahy), *A. halleri* (Ah) y *S. alfredii* (Sa). **(A).** Organización exón/intrón de los genes homólogos. Los rectángulos azules representan *UTR* (untranslated regions), los triángulos amarillos indican exones y las líneas negras discontinuas indican intrones. Los tamaños de los exones e intrones están representados según la escala en la parte inferior (Kb). **(B).** Patrón de Motivos conservados de proteínas homólogas identificados en MEME. Las coincidencias de los Motivos evidencian un *p-value* inferior a 0,00001. Las formas de un mismo color representan un Motivos conservado entre proteínas, mientras que las líneas negras indican regiones no conservadas. Los tamaños de las regiones conservadas y no conservadas están representadas según la escala en la parte inferior (aa). El árbol filogenético presentado se generó en *MEGAX* a partir de las secuencias de aminoácidos de longitud completa de las proteínas ZIP homólogas, utilizando el método Maximum Likelihood basado en el modelo de matriz *JTT* con valores de *Bootstrap* obtenidos de 1500 réplicas.

3.5. Análisis de metodologías para el establecimiento y regeneración in vitro de plantas de cacao

Los trabajos de investigación relacionados con el cultivo in vitro de tejidos para *T. cacao* fueron revisados uno a uno [79-88]. Los datos de los resultados principales obtenidos fueron ponderados en la **(Tabla S2)** y los trabajos de Li, Guillou, Ramirez Henao, Maximova, Traore y Gultinan [80, 83-86, 88] fueron tenidos en cuenta para proponer un protocolo de embriogénesis somática indirecta usando estaminodios de flores inmaduras. El protocolo incluye la obtención y el tratamiento del material vegetal, inducción de callos, embriogénesis somática primaria, embriogénesis somática secundaria y regeneración de plantas. Desde la obtención de callos hasta la obtención de plantas pueden pasar aproximadamente 10 a 12 meses y se describe detalladamente abajo en la sección de Metodología de edición (3.6.3).

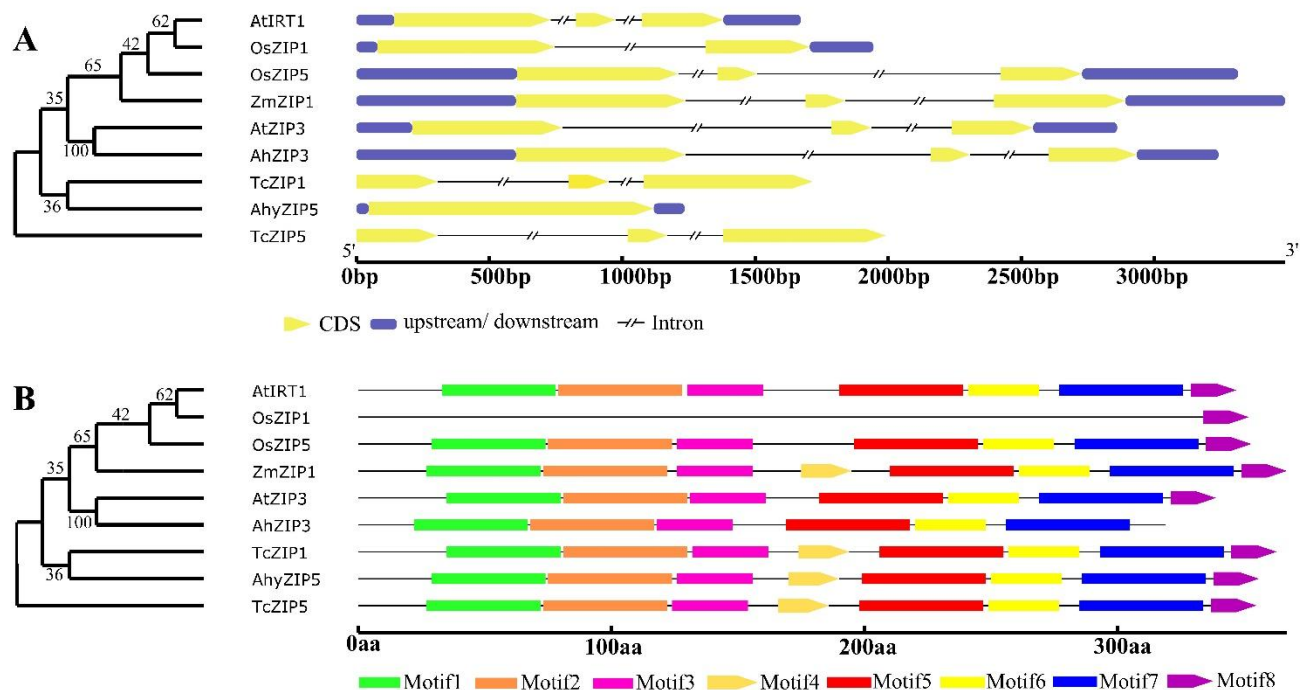


Figura 6. Estructura genética y arquitectura de Motivos conservados de miembros homólogos de la familia ZIP de *T. cacao* (Tc), *O. sativa* (Os), *A. thaliana* (At), *Z. mays* (Zm), *A. hypogaea* (Ah), *A. halleri* (Ah) y *S. alfredii* (Sa). **(A).** Organización exón/intrón de los genes homólogos. Los rectángulos azules representan *UTR* (untranslated regions), los triángulos amarillos indican exones y las líneas negras discontinuas indican intrones. Los tamaños de los exones e intrones están representados según la escala en la parte inferior (bp). **(B).** Patrón de Motivos conservados de proteínas homólogas identificados en MEME. Las coincidencias de los Motivos evidencian un *p-value* inferior a 0,00001. Las formas de un mismo color representan un Motivos conservado entre proteínas, mientras que las líneas negras indican regiones no conservadas. Los tamaños de las regiones conservadas y no conservadas están representadas según la escala en la parte inferior (aa). El árbol filogenético presentado se generó en *MEGAX* a partir de las secuencias de aminoácidos de longitud completa de las proteínas ZIP homólogas, utilizando el método Maximum Likelihood basado en el modelo de matriz *JTT* con valores de *Bootstrap* obtenidos de 1500 réplicas.

3.6. Estrategia de Edición genética

3.6.1. Diseño de ARN guía (*SgRNA*)

La secuencia del gen *TcNRAMP5*, incluyendo exones e intrones, fue sometida a la herramienta *CRISPR Site Finder* de *Geneious Prime*. Se obtuvieron 328 *CRISPR guides* que fueron filtrados de acuerdo a su ubicación, índice de actividad, cantidad de sitios *off-target* y porcentaje de especificidad. Así, se eligió como región *target* el exón V de *TcNRAMP5* ubicado en chr8:6825648..6830140. El *sgRNA* para esta región presenta un valor de especificidad del 99.62%, un índice de actividad de 0.636 (en un rango de 0 a 1) y un único sitio *off target* en chr8:6802469..6802491 con un porcentaje de emparejamiento del 1.91%. Adicionalmente, posee un sitio de restricción para las enzimas BsrBI, AccBSI y MbiI las cuales reconocen la secuencia CCGCTC (-3/-3) de la región *target* (**Figura 7**). Por otra parte, se usó la herramienta *Primers* de *Geneious Prime* para la generación de primers que permitan la identificación molecular del sitio de edición (**Tabla 2**). Los primers elegidos generan un amplicón de 450pb y cuenta con un %GC de 55%, un *Tm* de 60,2 °C y 59,8°C para el *Forward* y *Reverse* respectivamente.

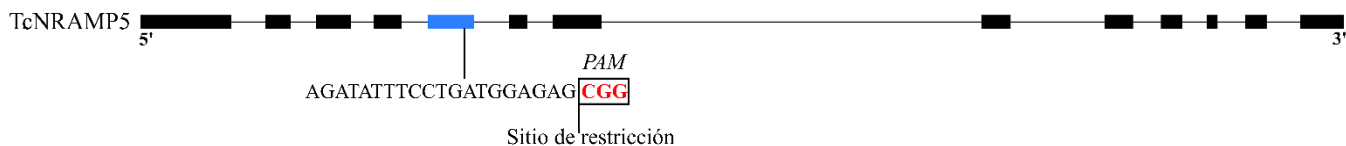


Figura 7. Diagrama esquemático de la estructura del gen *TcNRAMP5* con un sitio *target* CRISPR/Cas9. Los intrones se indican con una línea negra continua y los exones se representan con rectángulos negros y azul. La secuencia *target* de 20nt se muestra en la parte inferior del exón coloreado de azul, y la secuencia PAM (NGG) está resaltada en color rojo.

Primer	Secuencia	%GC	Tm	Hairpin Tm	Self Dimer Tm	Pair Dimer Tm
NR5F	GCTGCAGAGGCCGAAGTCATA	55	60,2	45,4	20,4	10,7
NR5R	GCTCTCGTCGTCTCCAGAAA	55	59,8	None	None	10,7

Tabla 2. Primers *Forward* y *Reverse* para la amplificación de la región de edición.

3.6.2. Elección del vector

Se seleccionó como vector el plásmido p55 para *Agrobacterium tumefaciens*. Este plásmido tiene un tamaño de 13314 pb, específico para eudicotiledoneas, el cual posee un gen marcador de selección de resistencia al herbicida *Bialafos* (Bar) en la región T (**Figura 8**) y el gen marcador de selección NeoR/KanR para conferir resistencia a Kanamicina y otros antibióticos en *A. tumefaciens*. Los promotores de expresión son el CaMV 35S para Bar-Cas9 y AtU6 para el *SgRNA*.

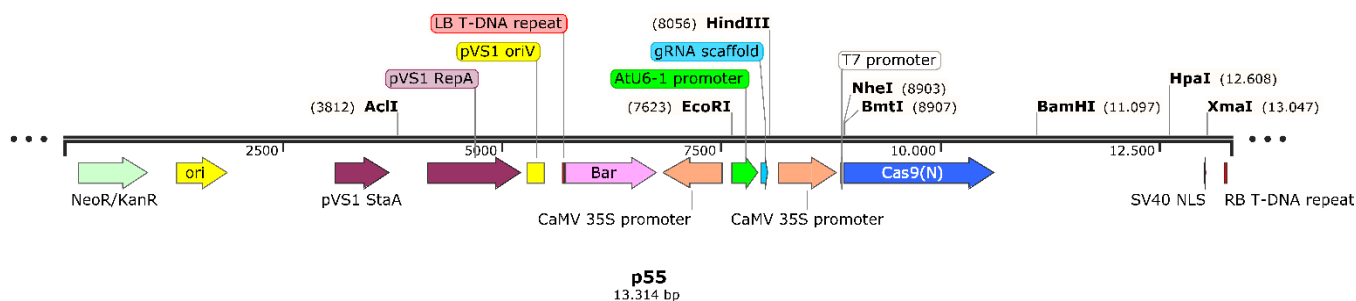


Figura 8. Estructura del constructo p55-CRISPR/Cas9 para transformación mediada por *Agrobacterium*.

3.6.3. Metodología de edición

De la literatura obtenida en el análisis de metadatos, se encontró que pocos trabajos sobre edición genética en *T. cacao* usando el sistema CRISPR/Cas9 han sido reportados. No obstante, se ha logrado la expresión transitoria del sistema en hojas mediada por *A. tumefaciens* para generar una delección en un gen relacionado a la depresión del sistema de respuesta inmune del cacao (*TcNPR3*) favoreciendo la resistencia a infecciones como la del hongo patógeno *Phytophthora tropicalis* [89-96].

Con lo reportado en literatura y demás metodologías consultadas, se estructuró un protocolo de transformación mediada por *A. tumefaciens* para embriones primarios de *T. cacao* provenientes de explantes florales (estaminodios). Para la transformación se propone usar la cepa AGL1 de *A. tumefaciens* transformada con el vector p55 que contiene el sistema CRISPR/Cas9, tal y como se describe en la metodología. *A. tumefaciens* AGL1 se ha probado con éxito en procesos de transformación en *T. cacao* y se encontró que, usando el antibiótico Moxalactam para eliminar de forma más efectiva los remanentes de *A. tumefaciens* después del proceso de infección, se puede aumentar la tasa de supervivencia de los explantes [80,92]. Posteriormente se seleccionan y multiplican los explantes transformados (por medio del proceso de embriogénesis secundaria) para poder así, ser finalmente convertidos en plantas que finalmente serán endurecidas y aclimatadas para ser llevadas al invernadero.

Por último, proponemos un protocolo de aislamiento de protoplastos foliares teniendo en cuenta resultados previos tanto en *T. cacao* como en *A. thaliana* [97-100]. Así, esperamos obtener líneas editadas genéticamente de *T. cacao* mediante la transfección mediada por polietilenglicol (PEG), libres de transgén.

Los protocolos propuestos son presentados a continuación.

Protocolo para embriogénesis somática indirecta y regeneración *in vitro* de plantas de *T. cacao*.

1. *Material vegetal y preparación del explantes*

1.1. Recolectar en campo, preferiblemente en la mañana (7:00-9:00am), flores inmaduras (4-5mm) de 2 a 3 semanas de edad.

NOTA: Flores cosechadas en la tarde o con un estado de desarrollo mayor pueden abrirse y disminuir la efectividad de la esterilización.

1.2. Almacenar los botones florales en tubos Falcon de 50mL con solución fría y estéril de sales basales DKW (2,66g/L). Conservar la cadena de frío hasta su uso.

NOTA: Si los botones florales no serán usados el mismo día, se debe realizar en campo un lavado con hipoclorito de sodio al 0,5% (v/v) antes de almacenar en frío, esto permite conservar las muestras máximo 3 días.

1.3. Decantar la solución DKW y sumergir los botones florales en solución 1% (v/v) de hipoclorito de sodio por 5 minutos bajo agitación constante en un recipiente de vidrio o plástico estéril.

1.4. Retirar el hipoclorito y realizar 3 lavados con agua destilada estéril y posteriormente almacenar en un nuevo tubo Falcon estéril.

1.5. Sumergir en solución de estreptomicina bajo agitación constante a 250ppm por 20 minutos.

1.6. Realizar un último lavado con agua destilada estéril y transferir entre 4 y 10 botones florales estériles a una caja petri estéril (100x15mm) y retirar los estaminodios cortando a 1/3 de la base del botón floral, haciendo uso de pinzas y escalpelo. Remover y descartar pétalos y sépalos.

NOTA: Evitar llevar exceso de agua en los botones florales y realizar el procedimiento lo más rápido posible para evitar disecación.

Ver guía fotográfica en: <https://plantscience.psu.edu/research/labs/gultinan/protocols/photographs-of-cacao-tissue-culture-protocols/dissecting-cacao-flowers-for-somatic-embryogenesis>

2. *Inducción de callos*

2.1. Sembrar los estaminodios estériles en cajas Petri (60x15mm) con 12mL de medio PCG (*Primary Callus Growth*) (**Tabla 3**) Incubar entre 25 a 30°C en oscuridad, por 2 semanas.

NOTA: Aproximadamente 15 a 25 estaminodios por caja. Asegurar el contacto con el medio (sin inmersión).

2.2. Pasadas las 2 semanas, cultivar en medio SCG (*Secondary Callus Growth*) (**Tabla 3**) Incubar a las mismas condiciones del numeral anterior.

NOTA: Después de 2 semanas bajo estas condiciones debe ser evidente la formación de callo para la finalización de este periodo de cultivo.

3. *Embriogénesis somática primaria (Maduración)*

3.1. Transferir los callos derivados de estaminodios a medio ED (*Embryo Development*) (**Tabla 3**). Mantener en oscuridad e incubar entre 25 a 30°C realizando subcultivos cada 2 semanas.

NOTA: Desde el primer subcultivo debe observarse un engrosamiento individual de los explantes y deben distinguirse 2 tipos de callo. TIPO 1: células alargadas y de color blanco fácilmente identificable bajo el estereoscopio (los embriones somáticos casi nunca se desarrollan a partir de este tipo de callo). TIPO 2: células

redondas y de color marrón claro a oscuro y friable. Adicionalmente, estos callos en medio ED pueden mantenerse para producción de embriones por un año, con un máximo de producción a las 24 semanas.

3.2. Una vez se observan embriones somáticos en estadio torpedo, se remueven del callo y son cultivados nuevamente en medio ED bajo las mismas condiciones del numeral anterior.

NOTA: Antes de obtener embriones en estadio torpedo, el embrión debe pasar por los estadios granular y corazón.

3.3. Embriones somáticos maduros con un eje embrionario completamente desarrollado son seleccionados para embriogénesis somática secundaria (numeral 4.) o para la regeneración de plantas (numeral 5.).

Ver guía fotográfica en: <https://plantscience.psu.edu/research/labs/gultinan/protocols/photographs-of-cacao-tissue-culture-protocols/stages-of-cacao-somatic-embryogenesis>

4. Transformación de explantes con *Agrobacterium tumefaciens* AGL1

El procedimiento de transformación se realiza con embriones somáticos primarios obtenidos del proceso de embriogénesis primaria descrita en el protocolo propuesto para establecimiento *in vitro*. Los embriones deben estar maduros, sanos, con cotiledones desarrollados y poseer una superficie brillante de color blanco, rosa o marfil opaco y sin tricomas visibles.

4.1. Cultivo e inducción de virulencia de *A. tumefaciens* AGL1.

4.1.1. Inocular 20 mL de medio 523 con *A. tumefaciens* AGL1 (con el vector) y agitar a 200rpm en un agitador orbital hasta obtener una O.D. de 1 a 420nm.

4.1.2. Centrifugar el medio con *A. tumefaciens* AGL1 a 800g por 20 minutos a 25°C.

4.1.3. Resuspender el pellet en medio ED líquido y sin hormonas, suplementado con 0,0001M de acetosiringona y 0,0013M de prolina. Ajustar la O.D. hasta 0,5 a 420nm. Inducir virulencia de *A. tumefaciens* AGL1 por agitación a 100rpm durante 5h a 25°C.

4.2. Infección de explantes

4.2.1. Cortar los cotiledones (provenientes de los embriones somáticos primarios) en trozos de 4mm².

4.2.2. Introducir entre 100 y 200 explantes (cotiledones) en tubos Falcon de 50mL con 20mL de solución de *A. tumefaciens* AGL1.

4.2.3. Sonicar la solución de *A. tumefaciens* con los explantes por 30s en un baño de agua.

4.2.4. Infectar por 10 minutos a 25°C y 50rpm.

4.2.5. Aspirar al vacío la solución de *A. tumefaciens* AGL1 y transferir los explantes a medio sólido SCG, cubrir los explantes con papel de filtro estéril e incubar por 48 horas de co-cultivo a 25°C en la oscuridad.

4.3. Selección de transformantes

4.3.1. Transferir los explantes a medio sólido SCG con 200 mg/L de Moxalactam y 5 o 3 mg/L de Bialaphos (G418). Cultivar los explantes en medio de selección durante 2 semanas en la oscuridad a 25° C.

4.3.2. Transferir y mantener los explantes en medio ED suplementado con 200mg/L de Moxalactam y seguir el protocolo de Embriogénesis secundaria.

4.3.3. Para generar embriones terciarios y cuaternarios, repetir el protocolo para Embriogénesis secundaria las veces que sea necesaria y seleccionar los transformantes como se describe en el numeral 3.1 y 3.2

5. Embriogénesis somática secundaria

➔ Obtención embriones en mayor cantidad y de mejor calidad.

5.1. Embriones somáticos primarios, con cotiledones desarrollados (1 a 4mm) de color amarillo claro o rosa, deben ser separados del hipocótilo y seccionados en 4 partes iguales usando un escalpelo y posteriormente llevarlos a medio MC2 mantener en oscuridad por 2 semanas entre 25 y 30°C. Ver composición completa en tabla 1.

NOTA: No pueden ser cotiledones jóvenes (vellosos, transparentes/blancos) o viejos (engrosados, con tricomas largos y visibles). Además, debe garantizarse el contacto de los explantes con el medio (sin inmersión). Después de 2 semanas bajo estas condiciones debe ser evidente la formación de callo para la finalización de este periodo de cultivo.

5.2. Transferir los explantes a medio ED y mantener bajo las mismas condiciones del numeral anterior y realizar subcultivos cada 2 semanas.

NOTA: Los embriones secundarios deben formarse entre 2 y 3 meses.

6. Conversión de embriones somáticos y establecimiento de plantas

6.1. Seleccionar embriones maduros de aproximadamente 2cm con cotiledones distintivos y un eje extendido. Sembrarlos verticalmente (ángulo 90°) en cajas petri (100x15mm) con medio PEC (*Primary Embryo Conversion*) (**Tabla 3**) e incubar entre 25 a 30°C manteniendo un fotoperiodo 12/12h (luz/oscuridad).

NOTA: Entre 5-10 embriones por caja y subcultivar mensualmente hasta obtener brotes con al menos 2 hojas desarrolladas (1cm de longitud) y raíces primarias.

6.2. Transferir brotes a recipientes Magenta GA3 con 143mL de medio RD (*Root Development and Maintenance*) para prevenir contaminación (**Tabla 3**).

NOTA: Para promover el crecimiento normal de los brotes se recomienda decapitar brotes anormales. Para esto, se corta con un escalpelo 5mm arriba del primer nodo (Asegurar dejar al menos un nodo para el crecimiento adventicio del brote). Adicionalmente, transferir a medio RD fresco cada 2 meses.

7. Aclimatación

7.1. Seleccionar plantas maduras, con al menos 4 hojas (más de 2cm c/u) y raíces sanas (más de 2cm de longitud) y sembrarlas en recipientes plásticos que contengan tierra autoclavada con buen drenaje (mezcla de arena y tierra). Revisar diariamente durante la primera semana y regar según sea necesario (no excederse). Además, es necesario adicionar fertilizante soluble en el agua (Puede usarse la solución de Hoagland).

NOTA: Antes de plantar se debe saturar con agua. Las plantas deben mantenerse en 100% humedad y 70% de sombra por 4 semanas.

7.2. Después de las 4 semanas, las plántulas deben pasar a tener una humedad entre el 60 y 65% con 50% de sombra.

NOTA: Durante este periodo de aclimatación las plantas deben alcanzar un tamaño de 5 a 10cm de altura.

7.3. Las plantas están listas para llevar a invernadero.

	Inducción de callos		Embriogénesis (Primaria)	Embriogénesis (Secundaria)	Conversión	
	PCG DKW	SCG WPM	ED DKW	MC2 MS	PEC DKW	RD DKW/2
Vitaminas (mg/L)						
Mioinositol	200,0	100,0	100,0	100,0	100,0	50,0
Ácido nicotínico	1,0	0,5	1,0	1,0	1,0	0,5
Tiamina-HCL	2,0	1,0	2,0	2,0	2,0	1,0
Piridoxina-HCL		0,5		0,5		
Ácido pantoténico		1,0				
Aminoácidos (mg/L)						
Glicina	2,0	2,0	2,0	2,0	2,187	1,0
L-Lisina					0,456	
L-Leucina					0,328	
L-Arginina					0,435	
L-Triptófano					0,511	
L-Glutamina	250,0					
Hormonas (mg/L)						
2,4,5 T			1,0	1,0		
2,4D	2,0	2,0				
ANA						
Kinetina		0,3				
Thidiazuron	0,005					
Isopentenil Adenina	50,0					
Adenina				0,25		
Agua de coco (%v/v)		5%				
Otros (g/L)						
Glucosa	20,0	20,0	1,0		20,0	10,0
Fitagel	2,0	2,2	2,0	2,0	1,75	1,75
Sacarosa		20,0	30,0	30,0	10,0	5,0
KNO ₃					0,3	0,3
pH	5,6	5,6	5,7	5,7	5,8	5,8

Tabla 3. Composición de medios de cultivos necesarios para el establecimiento en cultivo *in vitro* de plantas de *T. cacao* por medio de la embriogénesis somática indirecta a partir de explantes florales inmaduros. La tabla relaciona cada proceso con los medios usados en estos, sus componentes y nivel de pH.

Protocolo aislamiento y transfección de protoplastos

1. Crecimiento de plantas

- 1.1. Crear plantas de *T. cacao* en invernadero con una mezcla de suelo óptima para la germinación de semillas, con un tamaño de partícula pequeño y uniforme o en una cámara de ambiente controlado con un fotoperiodo 12/12h (luz/oscuridad), entre 25 a 30°C y una humedad relativa entre 50 y 65%.

2. Aislamiento de protoplastos

- 2.1. Colectar de 0,5 a 1 gramo de hojas bien expandidas de aproximadamente 4cm de largo, provenientes de plantas (3 a 5 semanas) crecidas en invernadero o, de 3cm de largo pero de plántulas obtenidas de cultivo *in vitro*.

NOTA: La selección de hojas sanas y en la etapa de desarrollo adecuada es un factor muy importante en los experimentos con protoplastos. Hojas de plantas estresadas pueden presentar bajos porcentajes de eficiencia de

transfección. Dejar la planta en la oscuridad por 24h permite fijar menos carbono en la pared celular, lo que disminuye la cantidad de almidón el cual hace difícil el ajuste osmótico.

- 2.2. Incubar las hojas en agua destilada estéril por 30 minutos. Pasados los 30 minutos y usando un escalpelo, cortar sobre una superficie de papel blanco estéril, tiras de hojas de forma transversal de 0,5 a 1mm de grosor.

NOTA: Cambiar la cuchilla después de 4 a 5 hojas para evitar aplastar el tejido en los sitios de corte. Una muestra clara de maltrato del tejido será evidente en la superficie de corte como una mancha jugosa y de color verde oscuro. Realizar rápido para evitar disecación del tejido.

- 2.3. Transferir las tiras de 4 a 5 hojas a una caja petri (100mmx15mm) con 5 a 10mL de solución enzimática, se recomienda hacerlo de forma suave y rápida (**Tabla 4**). Incubar en oscuridad a temperatura ambiente entre 3 y 4 horas y mantener en agitación constante a 50rpm.

NOTA: Sumergir completamente con la ayuda de pinzas. La solución debe tornarse verde, indicando la liberación de protoplastos, sin embargo, confirmar bajo el microscopio.

- 2.4. Antes de filtrar, diluir la solución enzimática con los protoplastos en un volumen igual de solución W5 (**Tabla 4**).
- 2.5. Usar una malla de nylon de 75µm previamente sumergida en etanol al 95%. Lavar la malla con agua destilada estéril y posteriormente humedecer con solución W5. Finalmente filtrar la solución enzimática con los protoplastos en un tubo falcon de 15 o 50mL según el volumen utilizado.
- 2.6. En una cámara de Neubauer contar protoplastos y estimar concentración por mililitro.
- 2.7. Centrifugar a 100 o 200g por 2 o 3 minutos para sedimentar los protoplastos. Retirar el sobrenadante haciendo uso de una micropipeta.
- 2.8. Resuspender el pellet (protoplastos) en solución W5 hasta obtener una concentración de 2×10^6 protoplastos/mL. Dejar reposar en hielo por 30 minutos. Retirar del hielo y dejar asentar los protoplastos por 15 a 25 minutos.
- 2.9. Resuspender en solución MMG (**Tabla 4**) con el mismo volumen usado en el numeral anterior. Contar nuevamente protoplastos en la cámara de Neubauer y obtener concentración por mililitro. Diluir hasta obtener una concentración de 2×10^5 .

3. *Transfección del complejo RNP usando PEG-calcio*

- 3.1. Adicionar 10µl del complejo ribonucleoproteína (RNP) a un tubo de microcentrífuga de 2ml.
- 3.2. Adicionar 100µl de solución de protoplastos (2×10^5) y mezclar cuidadosamente.
- 3.3. Adicionar 110µl de solución PEG (10 a 20%) y mezclar nuevamente. Incubar a temperatura ambiente por 15 minutos.
- 3.4. Diluir la mezcla con 400 a 440µl de solución W5 a temperatura ambiente e invertir el tubo suavemente para homogeneizar y detener la transfección.
- 3.5. Centrifugar a 100g durante 2 minutos a temperatura ambiente y remover el sobrenadante.
- 3.6. En una caja petri (60mmx15mm), resuspender los protoplastos en 1ml de solución WI (**Tabla 4**).
- 3.7. Incubar los protoplastos entre 20 a 25°C por el tiempo deseado.

Recomendaciones a tener en cuenta para la ejecución del protocolo

Para todas las soluciones, preparar primero con agua destilada estéril la solución de MES a la concentración indicada. Antes de agregar las enzimas, la solución de MES debe precalentarse a 70 ° C entre 3 a 5 minutos y filtrar la solución enzimática final a través de un filtro de jeringa estéril de 0,45µm en una caja petri (100x15mm) y

mantener en oscuridad. Preparar la solución PEG 1 hora antes de la transfección, puede almacenarse a temperatura ambiente por 5 días.

	Solución enzimática	W5	MMG	WI	PEG
Enzimas					
<i>Celullase R-10 (Onozuka)</i>	1,0%				
<i>Macerozyme R-10 (Onozuka)</i>	1,0%				
<i>Driselase</i>	0,5%				
Componentes (mM)					
MES	20,0	2,0	4,0	4,0	
Manitol	20,0		400,0	0,5	200,0
KCl	20,0	5,0		20,0	
CaCl ₂	10,0	125,0			100,0
BSA	0,1%				
NaCl		154,0			
MgCl ₂			15,0		
β-mercaptoetanol*	1,0 a 5,0				
PEG4000 en ddH ₂ O					10 a 20%
pH	5,7	5,7	5,7	5,7	

Tabla 4. Composición de soluciones para aislamiento y transfección de protoplastos de hojas de *T. cacao*. * Opcional.

4. Discusión

El cadmio es un metal pesado ampliamente distribuido en el mundo, liberado al medio ambiente principalmente por sistemas de calefacción, industrias metalúrgicas, incineradores de residuos, tráfico urbano, fábricas de cemento y fertilizantes de fosfato. Su alta presencia en suelos ha ocasionado que diversas especies de plantas de interés comercial, incluyendo *T. cacao*, acumulen este metal en diferentes órganos [101-103]. Esto ha llevado a que el cadmio sea catalogado como uno de los contaminantes que generan mayor preocupación a nivel mundial debido a su impacto negativo en la salud humana y la seguridad alimentaria. Por esta razón, comprender los mecanismos que rigen la acumulación de este metal es una necesidad para la generación de estrategias que permitan disminuir sus impactos negativos. Debido a que la abundancia y biodisponibilidad de micronutrientes metálicos puede ser limitante en algunos suelos, las plantas han desarrollado diferentes estrategias para su absorción [104]. De esta manera, desde un punto de vista genético, la absorción, transporte y acumulación están gobernadas por la acción de más de un gen principal (**Tabla 1**) [29]. No obstante, la identificación de genes clave, principalmente en el proceso de absorción, representa un objetivo prometedor para la disminución de la acumulación de cadmio en plantas mediante ingeniería genética.

Nuestros hallazgos dejan en evidencia que los miembros de las Familias de Proteínas Asociadas a cadmio FPAC (ABC, NRAMP, HMA, COPT, OPT, ZIP) han sido ampliamente reportados en *A. thaliana* y *O. sativa* (**Tabla 1**) por lo que estas fueron tenidas en cuenta para la búsqueda de un gen clave en el proceso de absorción radicular de cadmio en *T. cacao*. El análisis de IPP (**Figura S1 y S2**) arrojó que las FPAC halladas para ambas especies de plantas estaban significativamente agrupadas dentro de una posible red metabólica asociada con la absorción, transporte y acumulación de metales, entre ellos cadmio. Por su parte, la familia ABCC de transportadores, es impulsada por la hidrólisis de ATP y sus miembros pueden actuar tanto como exportadores como importadores de iones [105]. La familia NRAMP corresponde a una familia de transportadores de metales encontrados en diversos organismos, entre ellos bacterias, plantas y humanos [106]. La familia ZIP incluye transportadores para una amplia variedad de cationes incluyendo cadmio, hierro, manganeso y zinc, y por ello han sido estudiados con miras a la obtención de plantas hiperacumuladoras para la bioremediación de suelos contaminados [31, 107]. La subfamilia HMA, perteneciente a la familia P-Type ATPase, es impulsada por la hidrólisis de ATP para importar y exportar

diversos cationes incluyendo Ca, Zn, Co, Pb, Mn, Ag y Cd [73]. La familia COPT transporta cobre con mayor especificidad que las otras familias, y su sobreexpresión aumenta la sensibilidad a cobre; no obstante, son capaces de transportar otros metales catiónicos como Cd [31]. Por último, la familia OPT incluye proteínas encontradas tanto en eucariotas como procariotas, capaces de transportar péptidos y sideróforos principalmente de hierro [108].

Las raíces de las plantas representan el sitio principal para la absorción primaria de metales de los suelos a la cadena alimentaria. Este hecho implica que el cadmio primero ingresa a las células de las raíces, de manera oportunista, a través de transportadores iónicos de baja especificidad ubicados en la membrana plasmática, para ser posteriormente movilizado hacia las otras partes de la planta al ser cargado a los tejidos vasculares. Esto indica que los genes que codifican proteínas relacionadas con el proceso de absorción de cadmio deben estar principalmente expresados en la raíz en comparación con otros órganos o tejidos. La absorción de Cd mediada por transportadores iónicos, su traslocación y transporte hasta llegar a almacenarse en hojas, frutos y semillas, es un proceso que afecta de forma gradual el *fitness* de la planta, reduciendo su eficiencia en procesos como la fotosíntesis, la respiración, división y elongación celular o bloqueando ciertas rutas metabólicas [15,109].

El análisis de datos de expresión para las FPAC reveló que las proteínas transportadoras AtNRAMP1, AtNRAMP4, AtMTPA2, AtMTP10, AtHMA4, AtHMA5, AtHMA7, AtIRT1, AtZIP2 y AtZIP3 de *A. thaliana* y OsNRAMP5, OsNRAMP6, OsMTP7, OsHMA3, OsHMA5, OsCOPT1, OsZIP1, OsZIP5, OsZIP7, OsZIP9 y OsNRAMP5*, OsMTP7*, OsHMA4*, OsHMA5*, OsCOPT1*, OsZIP1* de *O. sativa* para los genotipos *Nipponbare* e *IR64* respectivamente, presentan una mayor expresión en la raíz comparadas con los otros miembros de sus familias y los otros 2 órganos analizados, esto sugiere que su expresión puede estar fuertemente ligada a la absorción de Cd u otros metales en células de raíz. Al comparar entre sí (interfamilia) la expresión en raíz de las proteínas resultantes del análisis anterior por medio del ANOVA de un solo factor y el método LSD de Fisher, se halló que AtNRAMP1, AtHMA7 y AtZIP2 de *A. thaliana* y OsCOPT1, OsZIP1, OsNRAMP5 (*Nipponbare*), OsCOPT1* y OsZIP1* (*IR64*) de *O. sativa*, fueron expresadas de forma significativamente mayor que las demás (95% de confianza y $\alpha=0,05$), sugiriendo un posible rol principal de estas proteínas en el proceso de absorción de iones metálicos.

Para tratar de soportar los resultados obtenidos hasta este punto, se hizo un análisis de expresión diferencial en raíz de muestras de *A. thaliana* y *O. sativa* (*IR64*), sometidas durante diferentes tiempos a deficiencia de hierro (Fe) o exceso de Cd para la primera y exceso de Cd o Pb para la segunda. Los resultados de este análisis revelaron que las proteínas transportadoras de *A. thaliana* AtNRAMP1, AtNRAMP4, AtHMA5, AtMTPA2, AtIRT1 son sensibles a condiciones de deficiencia de hierro (**Figura 4A**), mostrando una significativa regulación hacia arriba a excepción de AtZIP3 quien mostró ser regulada hacia abajo ($p\text{-value} \leq 0,05$). Aunque bajo el tratamiento con 100 μ M de ClCd_2 no se encontraron diferencias significativas en expresión al compararlas con las muestras control, probablemente por el bajo tiempo de exposición (6h), el mapa de calor (**Figura 4A**) muestra una posible regulación hacia arriba de AtNRAMP1 y una posible regulación hacia abajo de AtMTPA2 y AtHMA5 que hubiese podido ser más evidente con un tiempo de exposición mayor. Las proteínas restantes (AtNRAMP4, AtIRT1 y AtZIP3) parecieron no sufrir ningún tipo de regulación. Para *O. sativa* (*IR64*) los resultados mostraron regulación significativa hacia abajo de OsNRAMP5*, OsHMA3*, OsZIP1* y una regulación hacia arriba de OsZIP5* (**Figura 4B**) cuando fueron sometidas a 100 μ M de ClCd_2 , mientras que las proteínas sometidas al tratamiento con $\text{Pb}(\text{NO}_3)_2$ no mostraron ninguna regulación que fuese significativa ($p\text{-value} \leq 0,05$). Estos resultados sugieren que las proteínas reguladas guardan algún tipo de relación con la absorción radicular de cadmio.

Las proteínas que se mostraron reguladas, hacia arriba o hacia abajo según el análisis anterior, dieron como resultado 10 homólogos en *T. cacao* (**Datos S6**) después de ser sometidas al BLASTp, de las cuales la AtNRAMP1 y OsNRAMP5 arrojaron los mismos homólogos al igual que AtZIP3 y OsZIP5, sugiriendo una posible homología funcional conservada entre los representantes proteicos de las 3 especies. Para apoyar esta hipótesis se construyeron las filogenias usando las secuencias de aminoácidos de estas proteínas, adicionando los homólogos encontrados para *Z. mays*, *A. hypogaea*, *A. hallery* y *S. alfredii* (**Figura 5 y 6**). El análisis filogenético, la estructura genética y la arquitectura de Motivos conservados para los miembros homólogos de la familia ZIP, no soportaron ningún tipo de relacion suficiente que pudiera dar cuenta de una posible homología funcional entre la TcZIP5 y sus posibles

homólogos en *A. thaliana* y *O. sativa*, por lo que los resultados obtenidos en los análisis de expresión no podrían ser extrapolados a la proteína homóloga encontrada en *T. cacao*.

Para los homólogos de la familia NRAMP, por el contrario, se observó la formación de 3 grupos soportados por la filogenia resultante, las estructuras genéticas halladas y la arquitectura de Motivos conservados predicha. La TcNRAMP5 y la OsNRAMP5 mostraron mayor cercanía en la filogenia, haciendo parte de grupos hermanos cuya relación fue soportada por un valor de *Bootstrap* de 97, teniendo una estructura genética muy parecida y compartiendo 15 de los 16 Motivos predichos para la familia NRAMP. Adicionalmente, AhyNRAMP5 y ZmNRAMP5 se encuentran como clados hermanos de TcNRAMP5 y OsNRAMP5 respectivamente. Lo anterior sugiere una posible mayor cercanía funcional y estructural de la TcNRAMP5 con la OsNRAMP que con la AtNRAMP1 que se agrupó con la SaNRAMP5 (*Bootstrap* 74), aún estando *A. thaliana* relacionada filogenéticamente de forma más cercana con *T. cacao* que con *S. alfredii* [3].

Los resultados obtenidos concuerdan con los reportados en estudios relacionados con miembros de la familia NRAMP y la absorción de Cd [110,111]. Se reportó en un estudio que, OsNRAMP5 mostró ser más eficiente y rápida al momento de absorber Cd comparada con sus posibles homólogas en trigo y maíz [48]. En 2017, se obtuvieron resultados prometedores usando *CRISPR/Cas9* para la obtención de líneas *OsNRAMP5 knockout*, resultando en disminuciones significativas (98%) en las concentraciones de Cd en grano y sin afectar el rendimiento de la planta [112]. En 2019 y usando de nuevo el sistema *CRISPR/Cas9*, se produjeron mutantes *OsNRAMP5* que mostraron una mínima absorción de Cd en suelos contaminados con este metal, llegando a mostrar concentraciones en grano incluso por debajo de los valores permitidos por la normativa en China (0,06mg/kg versus 0,20mg/kg permitidos) y sin llegar a afectar el rendimiento de las plantas en 1 de 2 líneas mutantes [113]. Las similitudes encontradas en el presente estudio entre las secuencias y estructuras genéticas de OsNRAMP5 y TcNRAMP5, revelan un alto grado de homología estructural y posiblemente funcional, lo que postula a la TcNRAMP5 como un gen objeto de edición genética en *T. cacao*, en aras de la obtención de líneas editadas genéticamente para la disminución en la absorción y acumulación de Cd en plantas.

De esta manera, derivado de la ejecución de esta investigación, se estructuró una estrategia de edición mediada por *CRISPR/Cas9* dirigida a generar un *knockout* del gen *TcNRAMP5* (chr8:6825648..6830140). Para la estrategia se diseñó un *sgRNA* de 20nt con una secuencia *PAM* (NGG) (AGATATTTTCCTGATGGAGAGCGG), así como metodologías para cultivo *in vitro* y transformación mediada por *Agrobacterium* de embriones de cacao producidos a través embriogénesis somática indirecta a partir de estaminodios florales. También se propuso una metodología de edición genética para la transfección de protoplastos provenientes de hojas de cacao y mediada por PEG para el mismo gen. Los resultados que se esperan obtener de la ejecución de las metodologías propuestas podrían trazar una nueva ruta a seguir para conseguir una disminución significativa en las concentraciones de Cd presentes en las plantas de cacao y, por ende, en los productos que se derivan de sus frutos, asegurando que son seguros para el consumo humano y posibilitando el crecimiento continuo de la industria cacaotera nacional y latinoamericana.

5. Conclusiones

En este estudio un total de 8 Familias de Proteínas Asociadas a Cadmio (FPAC) fueron evaluadas en búsqueda de un gen clave asociado en el proceso de acumulación de cadmio, haciendo uso de herramientas computacionales. Se encontró que el gen *TcNRAMP5* de *T. cacao*, homólogo a *OsNRAMP5* de *O. sativa*, está fuertemente relacionado con el proceso de absorción de cadmio a nivel radicular. La generación de líneas *TcNRAMP5 knockout* podría generar plantas con bajo porcentaje de acumulación de cadmio como ya se demostró en *O. sativa*, por lo que se generó una estrategia de edición genética basada en *CRISPR/Cas9* dirigida a *TcNRAMP5* con la que se espera obtener el silenciamiento completo del gen en cuestión y una disminución considerable en el contenido de cadmio en plantas y granos cacao producido en Colombia y Latinoamérica, que se ajuste a los requerimientos internacionales, no represente ningún tipo de riesgo a la salud humana y la seguridad alimentaria.

6. Agradecimientos

Agradecemos a Catalina Restrepo Osorio M.Sc. en Biología (Departamento de Ciencias Biológicas, Universidad EAFIT), por su apoyo en la búsqueda de información y asesoramiento sobre cultivo *in vitro* de tejidos.

7. Declaración de conflictos de interés

Los autores de este trabajo declaran no tener ningún conflicto de interés. No hay financiación.

8. Referencias

- [1] K. Kubitzki, M.W. Chase, Introduction to Malvales, in: K. Kubitzki, C. Bayer (Eds.), Flowering Plants Dicotyledons, 5th ed., Springer, Verlag Berlin Heidelberg, 2003: pp. 12–16.
- [2] M.J. Guiltinan, Cacao, in: E.C. Pua, M.R. Davey (Eds.), Transgenic Crops V, Springer, Berlin, Heidelberg, New York, 2007: pp. 497–518.
- [3] The Catalogue of Life Partnership. APG IV: Angiosperm Phylogeny Group classification for the orders and families of flowering plants. Occurrence dataset. (2017). <http://www.GBIF.org> (consultado mayo 06, 2020).
- [4] J.C. Motamayor, A.M. Risterucci, P.A. Lopez, C.F. Ortiz, A. Moreno, C. Lanaud, Cacao domestication I: The origin of the cacao cultivated by the Mayas, *Heredity* (Edinb). 89 (2002) 380-386.
- [5] S. Zarrillo, N. Gaikwad, C. Lanaud, T. Powis, C. Viot, I. Lesur, O. Fouet, X. Argout, E. Guichoux, F. Salin, R.L. Solorzano, O. Bouchez, H. Vignes, P. Severts, J. Hurtado, A. Yepez, L. Grivetti, M. Blake, F. Valdez, The use and domestication of *Theobroma cacao* during the mid-Holocene in the upper Amazon, *Nat. Ecol. Evol.* 2 (2018) 1879–1888.
- [6] N. Ogata, A. Gómez-pompa, The Domestication and Distribution of *Theobroma cacao* L. in the Neotropics. *Chocolate in Measoamerica*, Florida Scholarsh. Online. 69 (2009) 1–2.
- [7] J. Cuatrecasas, Cacao and its Allies: A taxonomic revision of the genus *Theobroma*, *Natl. Herb.* 35 (1964) 379–607.
- [8] International Cocoa Organization (ICCO), Quarterly Bulletin of Cocoa Statistics, ICCO. 46 (2) (2020) 1.
- [9] V. Voora, S. Bermúdez, C. Larrea, Global Market Report: Cocoa, International Institute for Sustainable Development (IISD), 2019.
- [10] Cocoa and Chocolate Market Size, Share & COVID-19 Impact Analysis, By Type (Cocoa Butter, Cocoa Liquor, Cocoa Powder, Dark Chocolate, Milk Chocolate, White Chocolate and Filled Chocolate), By Application (Food and Beverages, Cosmetic, Pharmaceuticals, and Others), and Regional Forecast 2020-2027. Via Fortune Business Insights. (2020). <https://www.fortunebusinessinsights.com/amp/industry-reports/cocoa-and-chocolate-market-100075>. (consultado septiembre 28, 2020).
- [11] Fedecacao, Economía nacional. (2019). <http://www.fedecacao.com.co/portal/index.php/es/2015-02-12-17-20-59/nacionales> (consultado septiembre 28, 2020) .
- [12] P. Abbott, T. Benjamin, G. Burniske, M. Croft, M. Fenton, C. Kelly, M. Lundy, F. Rodriguez, M. Wilcox, An analysis of the supply chain of cacao in Colombia, United States Agency for International Development, USAID. Cali, CO (2018) 1-208.

- [13] F. Barraza, E. Schreck, T. Lévêque, G. Uzu, F. López, J. Ruales, J. Prunier, A. Marquet, L. Maurice, Cadmium bioaccumulation and gastric bioaccessibility in cacao: A field study in areas impacted by oil activities in Ecuador, *Environ. Pollut.* 229 (2017) 950–963.
- [14] A. Polle, A. Schützendübel, Heavy metal signalling in plants: linking cellular and organismic responses. in: H. Hirt, K. Shinozaki, (Eds.), *Plant Responses to Abiotic Stress*. Springer, Verlag Berlin Heidelberg. 2003: pp. 187–215.
- [15] G. DalCorso, S. Farinati, S. Maistri, A. Furini, How plants cope with cadmium: Staking all on metabolism and gene expression, *J. Integr. Plant Biol.* 50 (2008) 1268–1280.
- [16] G. Fatima, A.M. Raza, N. Hadi, N. Nigam, A.A. Mahdi, Cadmium in human diseases: It's more than just a mere metal, *Indian J. Clin. Biochem.* 34 (2019) 371–378.
- [17] J.J.V. Branca, G. Morucci, A. Pacini, Cadmium-induced neurotoxicity: Still much ado, *Neural Regen. Res.* 13 (2018) 1879–1882.
- [18] A.A. Pastorelli, R. Angeletti, G. Binato, M.B. Mariani, V. Cibin, S. Morelli, S. Ciardullo, P. Stacchini, Exposure to cadmium through Italian rice (*Oryza sativa* L.): Consumption and implications for human health, *J. Food Compos. Anal.* 69 (2018) 115–121.
- [19] European Food Safety Authority (EFSA), Cadmium dietary exposure in the European population, *EFSA J.* 10 (2012) 1–37.
- [20] European Union, Commission Regulation (EU) No 488/2014 of 12 May 2014 amending Regulation (EC) No 1881/2006 as regards maximum levels of cadmium in foodstuffs, *Off. J. Eur. Union.* 138 (2014) 75–79.
- [21] C.S. Jiménez Tobón, Estado legal mundial del cadmio en cacao (*Theobroma cacao*): fantasía o realidad, *Prod. + Limpia.* 10 (2015) 89–104.
- [22] J.A. Casteblanco, Técnicas de remediación de metales pesados con potencial aplicación en el cultivo de cacao, *La Granja.* 27 (2018) 21–35.
- [23] M.E. Londoño Rubio, AGROSAVIA lidera en Colombia la unificación de metodologías de análisis para la cuantificación de cadmio en cacao.(2019). <https://www.agrosavia.co/noticias/cadmio-en-cacao> (consultado septiembre 23, 2020).
- [24] D. Bertoldi, A. Barbero, F. Camin, A. Caligiani, R. Larcher, Multielemental fingerprinting and geographic traceability of *Theobroma cacao* beans and cocoa products, *Food Control.* 65 (2016) 46–53.
- [25] K.L.M. Zug, H.A. Huamaní Yupanqui, F. Meyberg, J.S. Cierjacks, A. Cierjacks, Cadmium accumulation in peruvian cacao (*Theobroma cacao* L.) and opportunities for mitigation, *Water. Air. Soil Pollut.* 230 (2019) 1–18.
- [26] A. Echeverry, H. Reyes, Determinación de la concentración de cadmio en un chocolate colombiano con 65% de cacao y chocolates extranjeros con diferentes porcentajes de cacao, *Entre Cienc. e Ing.* 19 (2016) 22–32.
- [27] D. Argüello, E. Chavez, F. Lauryssen, R. Vanderschueren, E. Smolders, D. Montalvo, Science of the total environment soil properties and agronomic factors affecting cadmium concentrations in cacao beans: A nationwide survey in Ecuador, *Sci. Total Environ.* 649 (2019) 120–127.
- [28] A. Meter, R. Atkinson, B. Laliberte, Cadmium in Cacao from Latin America and the Caribbean: A review of research and potential mitigation solutions, *Biodiversity International*, Rome, 2019.
- [29] V. Bert, P. Meerts, P. Saumitou-Laprade, P. Salis, W. Gruber, N. Verbruggen, Genetic basis of Cd tolerance and hyperaccumulation in *Arabidopsis halleri*, *Plant Soil.* 249 (2003) 9–18.

- [30] D. Ueno, N. Yamaji, I. Kono, C.F. Huang, T. Ando, M. Yano, J.F. Ma, Gene limiting cadmium accumulation in rice, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 16500–16505.
- [31] R. Chaffai, H. Koyama, Heavy metal tolerance in *Arabidopsis thaliana*, in: J.C. Kader, M. Delseny (Eds.), *Advances in botanical research*, Elsevier Ltd., 2011: 1–49.
- [32] Y. Song, L. Jin, X. Wang, Cadmium absorption and transportation pathways in plants, *Int. J. Phytoremediation*. 19 (2017) 133–141.
- [33] A. Gravot, P. Richaud, P. Auroy, L. Taconnat, J. Renou, A. Pugin, D. Wendehenne, Nitric oxide contributes to cadmium toxicity in *Arabidopsis* by promoting cadmium accumulation in roots and by up-regulating genes related to iron uptake, *Plant Physiol.* 149 (2009) 1302–1315.
- [34] S. Clemens, Molecular mechanisms of plant metal tolerance and homeostasis. *Planta*. 412 (2001) 475–486.
- [35] Y.O. Korshunova, D. Eide, W.G. Clark, M. Lou Guerinot, B. Himadri, The IRT1 protein from *Arabidopsis thaliana* is a metal transporter with a broad substrate range, *Plant Mol. Biol.* 40 (1999) 37–44.
- [36] E.L. Connolly, J.P. Fett, M. Lou Guerinot, Expression of the IRT1 metal transporter is controlled by metals at the levels of transcript and protein accumulation, *Plant Cell*. 14 (2002) 1347–1357.
- [37] C. Chen, Q. Cao, Q. Jiang, J. Li, R. Yu, G. Shi, Comparative transcriptome analysis reveals gene network regulating cadmium uptake and translocation in peanut roots under iron deficiency, *BMC Plant Biol.* 19 (2019) 1–14.
- [38] R. Yu, D. Li, X. Du, S. Xia, C. Liu, G. Shi, Comparative transcriptome analysis reveals key cadmium transport-related genes in roots of two pak choi (*Brassica rapa* L. ssp. *chinensis*) cultivars, *BMC Genomics*. 18 (2017) 1–14.
- [39] H. Nakanishi, I. Ogawa, Y. Ishimaru, S. Mori, K. Nishizawa, Iron deficiency enhances cadmium uptake and translocation mediated by the Fe²⁺ transporters OsIRT1 and OsIRT2 in rice, *Soil Sci. Plant Nutr.* 52 (2006) 464–469.
- [40] O.A. Kulaeva, V.E. Tsyganov, Molecular-genetic basis of cadmium tolerance and accumulation in higher plants, *Russ. J. Genet. Appl. Res.* 1 (2011) 349–360.
- [41] A. Liu, Z. Zhou, Y. Yi, G. Chen, Transcriptome analysis reveals the roles of stem nodes in cadmium transport to rice grain, *BMC Genomics*. 21 (2020) 1–16.
- [42] T. Yoshihara, H. Hodoshima, Y. Miyano, K. Shoji, H. Shimada, F. Goto, Cadmium inducible Fe deficiency responses observed from macro and molecular views in tobacco plants, *Plant Cell Rep.* 25 (2006) 365–373.
- [43] N.S. Pence, P.B. Larsen, S.D. Ebbs, D.L.D. Letham, M.M. Lasat, D.F. Garvin, D. Eide, L. V Kochian, The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*, *Natl. Acad. Sci.* 97 (2000) 4956–4960.
- [44] C. Chen, Q. Cao, Q. Jiang, J. Li, R. Yu, G. Shi, Comparative transcriptome analysis reveals gene network regulating cadmium uptake and translocation in peanut roots under iron deficiency, *BMC Plant Biol.* 19 (2019) 1–14.
- [45] S. Thomine, R. Wang, J.M. Ward, N.M. Crawford, J.I. Schroeder, Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to Nramp genes, *Natl. Acad. Sci.* 97 (2000) 4991–4996.
- [46] Y. Nevo, N. Nelson, The NRAMP family of metal-ion transporters, *Biochim. Biophys. Acta*. 1763 (2006) 609–620.

- [47] R. Nazar, N. Iqbal, A. Masood, M.I.R. Khan, S. Syeed, N.A. Khan, Cadmium toxicity in plants and role of mineral nutrients in its alleviation, *Am. J. Plant Sci.* 3 (2012) 1476–1489.
- [48] L.E. Williams, J.K. Pittman, J.L. Hall, Emerging mechanisms for heavy metal transport in plants, *Biochim. Biophys. Acta.* 1465 (2000) 104–126.
- [49] J. Zhang, E. Martinoia, Y. Lee, Vacuolar transporters for cadmium and arsenic in plants and their applications in phytoremediation and crop development, *Plant Cell Physiol.* 59 (2018) 1317–1325.
- [50] F.Q. Sui, J.-D. Chang, Z. Tang, W.-J. Liu, X.-Y. Huang, F.-J. Zhao, Nramp5 expression and functionality likely explain higher cadmium uptake in rice than in wheat and maize, *Plant Soil.* 433 (2018) 377–389.
- [51] A. Sasaki, N. Yamaji, K. Yokosho, J. Feng Ma, Nramp5 is a major transporter responsible for manganese and cadmium uptake in rice, *Plant Cell.* 24 (2012) 2155–2167.
- [52] Y. Ishimaru, R. Takahashi, K. Bashir, H. Shimo, T. Senoura, K. Sugimoto, K. Ono, M. Yano, S. Ishikawa, T. Arao, H. Nakanishi, N.K. Nishizawa, Characterizing the role of rice NRAMP5 in manganese, iron and cadmium transport, *Sci. Rep.* 2 (2012) 1–8.
- [53] M. Courbot, G. Willems, P. Motte, S. Arvidsson, N. Roosens, P. Saumitou-laprade, N. Verbruggen, A major quantitative trait locus for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal ATPase, *Plant Physiol.* 144 (2007) 1052–1065.
- [54] R. Takahashi, K. Bashir, Y. Ishimaru, N.K. Nishizawa, H. Nakanishi, The role of heavy-metal ATPases, HMAs, in zinc and cadmium transport in rice, *Plant Signal. Behav.* 7 (2012) 1605–1607.
- [55] D.G. Mendoza-Cózatl, T.O. Jobe, F. Hauser, J.I. Schroeder, Long-distance transport, vacuolar sequestration, tolerance, and transcriptional responses induced by cadmium and arsenic, *Curr. Opin. Plant Biol.* 14 (2011) 554–562.
- [56] C. Bernard, N. Roosens, P. Czernic, M. Lebrun, N. Verbruggen, A novel CPx-ATPase from the cadmium hyperaccumulator *Thlaspi caerulescens*, *Fed. Eur. Biochem. Soc. Lett.* 569 (2004) 140–148.
- [57] F. Verret, A. Gravot, P. Auroy, N. Leonhardt, P. David, L. Nussaume, A. Vavasseur, P. Richaud, Overexpression of AtHMA4 enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance, *FEBS Lett.* 576 (2004) 306–312.
- [58] H. Liu, H. Zhao, L. Wu, A. Liu, F. Zhao, W. Xu, Heavy metal ATPase 3 (HMA3) confers cadmium hypertolerance on the cadmium/zinc hyperaccumulator *Sedum plumbizincicola*, *New Phytol.* 3 (2017) 687–698.
- [59] H. Miyadate, S. Adachi, A. Hiraizumi, K. Tezuka, N. Nakazawa, T. Kawamoto, K. Katou, I. Kodama, K. Sakurai, H. Takahashi, N. Satoh-Nagasawa, A. Watanabe, T. Fujimura, H. Akagi, OsHMA3, a P1B-type of ATPase affects root-to-shoot cadmium translocation in rice by mediating efflux into vacuoles, *New Phytol.* 189 (2011) 190–199.
- [60] D. Kim, L. Bovet, M. Maeshima, E. Martinoia, Y. Lee, The ABC transporter AtPDR8 is a cadmium extrusion pump conferring heavy metal resistance, *Plant J.* 50 (2007) 207–218.
- [61] Y. Kobae, T. Sekino, H. Yoshioka, T. Nakagawa, E. Martinoia, M. Maeshima, Loss of AtPDR8, a plasma membrane ABC transporter of *Arabidopsis thaliana*, causes hypersensitive cell death upon pathogen infection, *Plant Cell Physiol.* 47 (2006) 309–318.
- [62] M. Stein, J. Dittgen, C. Sa, B. Hou, A. Molina, P. Schulze-lefert, V. Lipka, S. Somerville, Arabidopsis PEN3 / PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter by direct penetration, *Plant Cell.* 18 (2006) 731–746.

- [63] Z. Hu, Y. Zhang, Y. He, Q. Cao, T. Zhang, L. Lou, Q. Cai, Full-length transcriptome assembly of Italian ryegrass root integrated with RNA-Seq to identify genes in response to plant cadmium stress, *Int. J. Mol. Sci.* 21 (2020) 1067.
- [64] J. Park, W.Y. Song, D. Ko, Y. Eom, T.H. Hansen, M. Schiller, T.G. Lee, E. Martinoia, Y. Lee, The phytochelatin transporters AtABCC1 and AtABCC2 mediate tolerance to cadmium and mercury, *Plant J.* 69 (2012) 278–288.
- [65] H. Küpper, B. Leitenmaier, Cadmium-accumulating plants, in: A. Sigel, H. Sigel, R. KO Sigel (Eds.), *Cadmium From Toxicity to Essentiality. Metal Ions in Life Sciences*, Springer, Dordrecht, 2013: pp. 373–393.
- [66] V. Korenkov, K. Hirschi, J.D. Crutchfield, G. Wagner, Enhancing tonoplast Cd/H antiport activity increases Cd, Zn, and Mn tolerance, and impacts root/shoot Cd partitioning in *Nicotiana tabacum* L., *Planta.* 226 (2007) 1379–1387.
- [67] V. Korenkov, B. King, K. Hirschi, G.J. Wagner, Root-selective expression of AtCAX4 and AtCAX2 results in reduced lamina cadmium in field-grown *Nicotiana*, *Plant Biotechnol. J.* 7 (2009) 219–226.
- [68] W. Song, E. Martinoia, J. Lee, D. Kim, D. Kim, E. Vogt, D. Shim, K.S. Choi, I. Hwang, Y. Lee, A novel family of Cys-rich membrane proteins mediates cadmium resistance in Arabidopsis, *Plant Physiol.* 135 (2004) 1027–1039.
- [69] R. Takahashi, Y. Ishimaru, T. Senoura, H. Shimo, S. Ishikawa, T. Arao, H. Nakanishi, N. Nishizawa, The OsNRAMP1 iron transporter is involved in Cd accumulation in rice, *J Exp Bot* 62 (2011) 4843–4850.
- [70] X. Zheng, L. Chen, X. Li, Arabidopsis and rice showed a distinct pattern in ZIPs genes expression profile in response to Cd stress, *Bot. Stud.* 59 (2018) 1–10.
- [71] M.L. Guerinot, The ZIP family of metal transporters, *Biochim. Biophys. Acta.* 1465 (2000) 190–198.
- [72] J.L. Gustin, M.J. Zanis, D.E. Salt, Structure and evolution of the plant Cation Diffusion Facilitator family of ion transporters, *BMC Evol. Biol.* 11 (2011) 76.
- [73] K.B. Axelsen, M.G. Palmgren, Inventory of the superfamily of P-type Ion pumps in Arabidopsis, *Plant Physiol.* 126 (2001) 696–706.
- [74] J.K. Pittman, K.D. Hirschi, CAX-ing a wide net: Cation/H(+) transporters in metal remediation and abiotic stress signalling, *Plant Biol. (Stuttg).* 18 (2016) 741–749.
- [75] K.M. Gomolplitinant, M.H. Saier, Evolution of the oligopeptide transporter family, *J. Membr. Biol.* 240 (2011) 89–110.
- [76] M.G. Stacey, H. Osawa, A. Patel, W. Gassmann, G. Stacey, Expression analyses of Arabidopsis oligopeptide transporters during seed germination, vegetative growth and reproduction, *Planta.* 223 (2006) 291–305.
- [77] P. Verrier¹, D. Bird, B. Burla, E. Dassa, C. Forestier, M. Geisler, M. Klein, Ü. Kolukisaoglu, Y. Lee, E. Martinoia, A. Murphy, P.A. Rea, L. Samuels, B. Schulz, E.J. Spalding, K. Yazaki, F.L. Theodoulou, Plant ABC proteins: a unified nomenclature and updated inventory, *Trends in Plant Science.* 13 (2008) 151–159.
- [78] S. Thomine, J.I. Srschroeder, Plant Metal Transporters with Homology to Proteins of the NRAMP Family, in: M. Cellier, P. Gros (Eds.), *The NRAMP Family*, Springer, Verlag Berlin Heidelberg, 2004: pp. 113–121.
- [79] W.P. Flynn, L.J. Glicenstein, P.J. Fritz, *Theobroma cacao* L.: an axillary bud *in vitro* propagation procedure, *Plant Cell. Tissue Organ Cult.* 20 (1990) 111–117.
- [80] Z. Li, A. Traore, S. Maximova, M.J. Guiltinan, Somatic embryogenesis and plant regeneration from floral explants of cacao (*Theobroma cacao* L.) using Thidiazuron, *Vitr. Cell. Dev. Biol. Plant.* (1998) 293–299.

- [81] C.H. Tsai, J.E. Kinsella, Initiation and growth of callus and cell suspensions of *Theobroma cacao* L, Ann. Bot. 48 (1981) 549–558.
- [82] C.L. Tan, D.B. Furtek, Development of an *in vitro* regeneration system for *Theobroma cacao* from mature tissues, Plant Sci. 164 (2003) 407–412.
- [83] C. Guillou, A. Fillodeau, E. Brulard, D. Breton, S.D.F. Maraschin, D. Verdier, M. Simon, J. Ducos, Indirect somatic embryogenesis of *Theobroma cacao* L. in liquid medium and improvement of embryo-to-plantlet conversion rate, 54 (2018) 377–391.
- [84] A.M. Ramírez Henao, T. Vasquez, M. Ospina Osorio, L. Garcés Atehortúa, A.I. Urrea Trujillo, Scientia Horticulturae evaluation of the potential of regeneration of different Colombian and commercial genotypes of cocoa (*Theobroma cacao* L.) via somatic embryogenesis, Sci. Hortic. (Amsterdam). 229 (2018) 148–156.
- [85] S.N. Maximova, L. Alemanno, A.N.N. Young, N. Ferriere, A. Traore, M.J. Guiltinan, Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L., Vitro. Cell. Dev. Biol. - Plant. 38 (2002) 252–259.
- [86] A. Traore, S.N. Maximova, M.J. Guiltinan, Micropropagation of *Theobroma cacao* L. using somatic embryo-derived plants, Vitro. Cell. Dev. Biol. - Plant. 39 (2003) 332–337.
- [87] S.N. Maximova, A. Young, S. Pishak, C. Miller, A. Traore, M.J. Guiltinan, Integrated system for propagation of *Theobroma cacao*, in: S. Jain, P. Gupta (Eds.), Protoc. Somat. Embryog. Woody Plants, Springer, Dordrecht, The Netherlands, 2005: pp. 209–229.
- [88] M.J. Guiltinan, S.N. Maximova, Integrated system for vegetative propagation of cacao V2.1. (2010). http://www.personal.psu.edu/mjg9/Guiltinan_Lab_Website/Protocols_files/ (consultado octubre 16, 2020).
- [89] S. Maximova, C. Miller, G. Antúnez De Mayolo, S. Pishak, A. Young, M.J. Guiltinan, Stable transformation of *Theobroma cacao* L. and influence of matrix attachment regions on GFP expression, Plant Cell Rep. 21 (2003) 872–883.
- [90] E.E. Helliwell, J. Vega-Arreguín, Z. Shi, B. Bailey, S. Xiao, S.N. Maximova, B.M. Tyler, M.J. Guiltinan, Enhanced resistance in *Theobroma cacao* against oomycete and fungal pathogens by secretion of phosphatidylinositol-3-phosphate-binding proteins, Plant Biotechnol. J. 14 (2016) 875–886.
- [91] A.S. Fister, L. Landherr, S.N. Maximova, M.J. Guiltinan, Transient expression of *CRISPR/Cas9* machinery targeting TcNPR3 enhances defense response in *Theobroma cacao*, Front. Plant Sci. 9 (2018) 1–15.
- [92] G. Antúnez De Mayolo, S.N. Maximova, S. Pishak, M.J. Guiltinan, Moxalactam as a counter-selection antibiotic for *Agrobacterium*-mediated transformation and its positive effects on *Theobroma cacao* somatic embryogenesis, Plant Sci. 164 (2003) 607–615.
- [93] S.N. Maximova, J.P. Marelli, A. Young, S. Pishak, J.A. Verica, M.J. Guiltinan, Over-expression of a cacao class I chitinase gene in *Theobroma cacao* L. enhances resistance against the pathogen, *Colletotrichum gloeosporioides*, Planta. 224 (2006) 740–749.
- [94] A.S. Fister, Z. Shi, Y. Zhang, E.E. Helliwell, S.N. Maximova, M.J. Guiltinan, Protocol: Transient expression system for functional genomics in the tropical tree *Theobroma cacao* L., Plant Methods. 12 (2016) 1–13.
- [95] Z. Shi, Y. Zhang, S.N. Maximova, M.J. Guiltinan, TcNPR3 from *Theobroma cacao* functions as a repressor of the pathogen defense response, BMC Plant Biol. 13 (2013) 1–12.
- [96] S.L. Florez, R.L. Erwin, S.N. Maximova, M.J. Guiltinan, W.R. Curtis, Enhanced somatic embryogenesis in *Theobroma cacao* using the homologous BABY BOOM transcription factor, BMC Plant Biol. 15 (2015).

- [97] H.A. Collin, R.B. Muse, S. Isaac, K. Hardwick, Isolation and properties of protoplasts of *Theobroma cacao* L. in: K.J. Puite, J.J.M. Dons, H.J. Huizing, A.J. Kool, M. Koornneef, F.A. Krens (Eds.), Progress in Plant Protoplast Research. Current Plant Science and Biotechnology in Agriculture, Springer, Dordrecht, 1988: pp.117–118.
- [98] S. Yoo, Y. Cho, J. Sheen, Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis, Nat Protoc. 2 (2007) 1565–1572.
- [99] K. Kanchanapoom, Cell suspension, isolation and culture of cacao (*Theobroma cacao* L.) protoplasts, J. Sci. Biotech. 21 (1999) 1-18.
- [100] F. Wu, S. Shen, L. Lee, S.H. Lee, M.T. Chan, C.S. Lin, Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method, Plant Methods. 5 (2009) 1-10.
- [101] M. Wang, J. Zou, X. Duan, W. Jiang, D. Liu, Cadmium accumulation and its effects on metal uptake in maize (*Zea mays* L.), Biosource Technol. 98 (2007) 82–88.
- [102] H.A. Huamaní-Yupanqui, M.Á. Huauya-Rojas, L.G. Mansilla-Minaya, N. Florida-Rofner, G.M. Neira-Trujillo, Presence of heavy metals in organic cacao (*Theobroma cacao* L.) crop, Acta Agronómica. 61 (2012) 309–314.
- [103] Z. Ramezani, M.J. Khodayar, Z. Nazari, R. Hedayatifar, A. Gilani, Cadmium and lead contents of rice (*Oryza sativa*) grown in Khuzestan, southwest of Iran, Jetashapir, J. Heal. Res. 5 (2014) 65–71.
- [104] H. Marschner, V. Romheld, Strategies of plants for acquisition of iron. Plant Soil. 165 (1994) 261–274.
- [105] J. Kang, J. Park, H. Choi, B. Burla, T. Kretschmar, Y. Lee, E. Martinoia, Plant ABC Transporters. The arabidopsis book. 9 (2011) 1-25.
- [106] Y. Nevo, N. Nelson, The NRAMP family of metal-ion transporters. Biochim. Biophys. Acta. 7 (2006) 609–20.
- [107] U. Krämer, I.N. Talke, M. Hanikenne, Transition metal transport, FEBS Lett. 581 (2007) 2263–2272.
- [108] K.M. Gomolplitinant, M.H. Saier, Evolution of the Oligopeptide Transporter Family, J. Membrane Biol. 240 (2011) 89–110
- [109] H. Bashir, M.I. Qureshi, M.M. Ibrahim, M. Iqbal, Chloroplast and photosystems: Impact of cadmium and iron deficiency, Photosynthetica. 53 (2015) 321–335.
- [110] I. Ullah, Y. Wang, D.J. Eide, J.M. Dunwell, Evolution, and functional analysis of Natural Resistance-Associated Macrophage Proteins (NRAMPs) from *Theobroma cacao* and their role in cadmium accumulation, Sci. Rep. 8 (2018) 1–15.
- [111] R.E.T. Moore, I. Ullah, V.H. de Oliveira, S.J. Hammond, S. Strekopytov, M. Tibbett, J.M. Dunwell, M. Rehkämper, Cadmium isotope fractionation reveals genetic variation in Cd uptake and translocation by *Theobroma cacao* and role of natural resistance-associated macrophage protein 5 and heavy metal ATPase-family transporters, Hortic. Res. 7 (2020).
- [112] L. Tang, B. Mao, Y. Li, Q. Lv, L. Zhang, C. Chen, H. He, W. Wang, X. Zeng, Y. Shao, Y. Pan, Y. Hu, Y. Peng, X. Fu, H. Li, S. Xia, B. Zhao, Knockout of *OsNramp5* using the *CRISPR/Cas9* system produces low Cd-accumulating indica rice without compromising yield, Sci. Rep. 7 (2017) 1-12.
- [113] L. Songmei, J. Jie, L. Yang, M. Jun, X. Shouling, T. Yuanyuan, L. Youfa, S. Qingyao, H. Jianzhong, Characterization and evaluation of *OsLCT1* and *OsNramp5* mutants generated through *CRISPR/Cas9*-Mediated mutagenesis for breeding low Cd rice, Rice Sci. 26 (2019) 88–97.